Transfer Theorems and Asymptotic Distributional Results for m-ary Search Trees

JAMES ALLEN FILL[1]

Department of Applied Mathematics and Statistics

The Johns Hopkins University

jimfill@jhu.edu and http://www.ams.jhu.edu/~fill/

AND

NEVIN KAPUR[1 2]

Department of Computer Science

California Institute of Technology

nkapur@cs.caltech.edu and http://www.cs.caltech.edu/~nkapur/

ABSTRACT

We derive asymptotics of moments and identify limiting distributions, under the random permutation model on m-ary search trees, for functionals that satisfy recurrence relations of a simple additive form. Many important functionals including the space requirement, internal path length, and the so-called shape functional fall under this framework. The approach is based on establishing *transfer theorems* that link the order of growth of the input into a particular (deterministic) recurrence to the order of growth of the output. The transfer theorems are used in conjunction with the method of moments to establish limit laws. It is shown that (i) for small toll sequences (t_n) [roughly, $t_n = O(n^{1/2})$] we have asymptotic normality if $m \le 26$ and typically periodic behavior if $m \ge 27$; (ii) for moderate toll sequences [roughly, $t_n = \omega(n^{1/2})$ but $t_n = o(n)$] we have convergence to non-normal distributions if $m \le m_0$ (where $m_0 \ge 26$) and typically periodic behavior if $m \ge m_0 + 1$; and (iii) for large toll sequences [roughly, $t_n = \omega(n)$] we have convergence to non-normal distributions for all values of m.

AMS 2000 *subject classifications.* Primary 68W40; secondary 60F05, 68P10, 60E05.

Key words and phrases. Transfer theorems, m-ary search trees, additive functionals, random permutation model, limit distribution, Euler differential equation, indicial polynomial.

Date. Revised January 11, 2004.

[1]Research for both authors supported by NSF grant DMS–9803780 and DMS–0104167, and by The Johns Hopkins University's Acheson J. Duncan Fund for the Advancement of Research in Statistics.

[2]Research supported by NSF grant 0049092 and carried out primarily while this author was affiliated with what is now the Department of Applied Mathematics and Statistics at The Johns Hopkins University.

1 Background and notation

We start by giving a brief overview of search trees which are fundamental data structures in computer science used in searching and sorting. For integer $m \geq 2$, the m-ary search tree, or multiway tree, generalizes the binary search tree. The quantity m is called the *branching factor*. According to [17], search trees of branching factors higher than 2 were first suggested by Muntz and Uzgalis [21] "to solve internal memory problems with large quantities of data." For more background we refer the reader to [15, 16] and [17].

An *m-ary tree* is a rooted tree with at most m "children" for each *node (vertex)*, each child of a node being distinguished as one of m possible types. Recursively expressed, an m-ary tree either is empty or consists of a distinguished node (called the *root*) together with an ordered m-tuple of *subtrees*, each of which is an m-ary tree.

An *m-ary search tree* is an m-ary tree in which each node has the capacity to contain $m-1$ elements of some linearly ordered set, called the set of *keys*. In typical implementations of m-ary search trees, the keys at each node are stored in increasing order and at each node one has m pointers to the subtrees. By spreading the input data in m directions instead of only 2, as is the case for a binary search tree, one seeks to have shorter path lengths and thus quicker searches.

We consider the space of m-ary search trees on n keys, and assume that the keys are linearly ordered. Hence, without loss of generality, we can take the set of keys to be $[n] := \{1, 2, \dots, n\}$. We construct an m-ary search tree from a sequence s of n distinct keys in the following way:

(i) If $n < m$, then all the keys are stored in the root node in increasing order.

(ii) If $n \geq m$, then the first $m-1$ keys in the sequence are stored in the root in increasing order, and the remaining $n-(m-1)$ keys are stored in the subtrees subject to the condition that if $\sigma_1 < \sigma_2 < \cdots < \sigma_{m-1}$ denotes the ordered sequence of keys in the root, then the keys in the jth subtree are those that lie between σ_{j-1} and σ_j, where $\sigma_0 := 0$ and $\sigma_m := n+1$, sequenced as in s.

(iii) All the subtrees are m-ary search trees that satisfy conditions (i), (ii), and (iii).

It is our goal to study additive functionals (see Definition 1.1) defined on m-ary search trees. Such functionals represent the cost of divide-and-conquer algorithms, reflecting the inherent recursive nature of the algorithms.

Let T be an m-ary search tree. We use $|T|$ to denote the number of keys in T. Call a node *full* if it contains $m-1$ keys. For $1 \leq j \leq m$, let $L_j(T)$ denote the jth subtree pendant from the root of T. For a node x in T, write T_x for the subtree of T consisting of x and its descendants, with x as root. This notation is illustrated in Figure 1.

Definition 1.1. Fix $m \geq 2$. We will call a functional f on m-ary search trees an *additive tree functional* if it satisfies the recurrence

$$f(T) = \sum_{i=1}^{m} f(L_i(T)) + t_{|T|}, \tag{1.1}$$

for any tree T with $|T| \geq m-1$. Here $(t_n)_{n \geq m-1}$ is a given sequence, henceforth called the *toll sequence* or *toll function*.

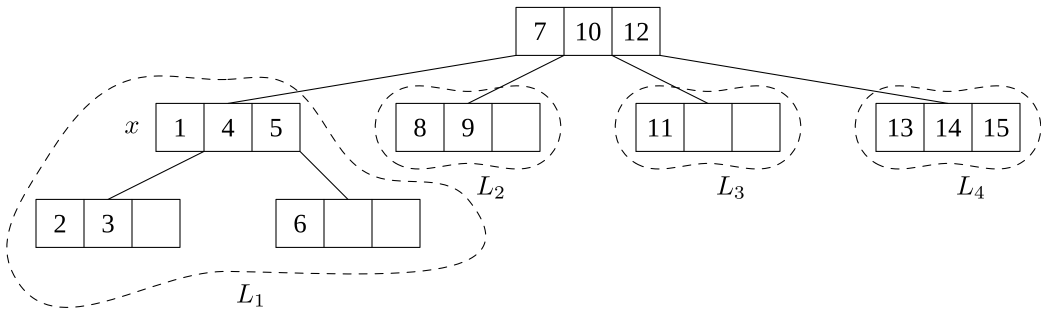

Figure 1: Example of notation for a quaternary tree T: here, $|T| = 15$ and $T_x = L_1$.

Note that the recurrence (1.1) does not make any reference to t_n for $0 \le n \le m-2$ nor specify $f(T)$ for $0 \le |T| \le m-2$.

Several interesting examples can be cast as additive functionals.

Example 1.2. If we specify $f(T)$ arbitrarily for $0 \le |T| \le m-2$ and take $t_n \equiv c$ for $n \ge m-1$, we obtain the "additive functional" framework of [17, §3.1]. (Our definition of an additive functional generalizes this notion.) In particular if we define $f(\emptyset) := 0$ and $f(T) := 1$ for the unique m-ary search tree T on n keys for $1 \le n \le m-2$ and let $t_n \equiv 1$ for $n \ge m-1$, then $f(T)$ counts the number of nodes in T and thus gives the *space requirement* functional discussed in [17, §3.4].

Example 1.3. If we define $f(T) := 0$ when $0 \le |T| \le m-2$ and $t_n := n-(m-1)$ for $n \ge m-1$ then f is the *internal path length* functional discussed in [17, §3.5]: $f(T)$ is the sum of all root-to-key distances in T.

Example 1.4. As described above, each permutation of $[n]$ gives rise to an m-ary search tree. Suppose we place the uniform distribution on such permutations. This induces a distribution on m-ary search trees called the *random permutation model*. Denote its probability mass function by Q. It is important to note that Q is *not* uniform, since different permutations can give rise to the same tree. For example, the permutations

$$(10, 7, 12, 4, 1, 8, 5, 6, 9, 14, 11, 2, 15, 13, 3)$$

and

$$(7, 10, 12, 1, 4, 8, 5, 6, 9, 14, 11, 2, 15, 13, 3)$$

both give rise to the quaternary search tree shown in Figure 1. Dobrow and Fill [8] noted that

$$Q(T) = \frac{1}{\prod_x \binom{|T_x|}{m-1}}, \tag{1.2}$$

where the product in (1.2) is over all full nodes in T. This functional is sometimes called the *shape functional* as it serves as a crude measure of the "shape" of the tree, with "full" trees (like the complete tree) achieving larger values of Q. For further discussions along these lines, consult [8] and [9]. If we define $f(T) := 0$ for $0 \le |T| \le m-2$ and $t_n := \ln \binom{n}{m-1}$ for $n \ge m-1$, then $f(T) = -\ln Q(T)$. In this work we will refer to

$-\ln Q$ (rather than Q) as the shape functional. It was our desire to understand the distribution of the shape functional under the random permutation model that led to this paper.

We study the distribution of the functional $f(T)$ when T is given the distribution Q described in Example 1.4. To do this, we derive asymptotics for the moments of $f(T)$ and then employ the method of moments.

Remark. In the sequel (without any loss of applicability) we will restrict attention to real-valued toll sequences.

Related work

Chern and Hwang [5] carried out the program of this paper for the space requirement (Example 1.2) and independently discovered Theorem 2.4 [5, Proposition 7]. Hwang and Neininger [14] analyzed a range of (not necessarily deterministic) toll functions for binary search trees. In this paper we treat m-ary search trees for any $m \geq 2$, and for simplicity we restrict attention to deterministic toll sequences. Another closely related paper is [6], where the authors consider variants of the `Quicksort` algorithm by allowing more general schemes of choosing the pivot. As in our Section 2.2, their paper treats Cauchy–Euler differential equations, and a generalization of our Theorem 2.4 is obtained [6, Theorem 1].

We obtain moment asymptotics for our additive functionals using the Asymptotic Transfer Theorem of Section 2.2. An alternative approach employs singularity analysis [12] of generating functions. A sketch of this approach in the case of binary search trees ($m = 2$) is presented in [10]. One small advantage of the present approach is that the conditions we impose on the toll sequence [for example, (2.11)] are milder than those required for the application of singularity analysis.

2 Overview: main results

For a given toll sequence, the distribution of $f(T)$ depends only on n. We let X_n denote a random variable whose distribution is that of $f(T)$ under the random permutation model on T. It is the distribution of X_n which is the main focus of this paper.

2.1 A common framework for all moments

Under the random permutation model the joint distribution of the subtree sizes $|L_1|, \ldots, |L_m|$ is uniform over all $\binom{n}{m-1}$ m-tuples of nonnegative integers that sum to $n-(m-1)$: see [17, Exercise 3.8]. We now apply the law of total expectation to compute $\mu_n(k) := \mathbf{E}\, X_n^k$ by conditioning on $(|L_1|, \ldots, |L_m|)$. Let $\sum_{\mathbf{j}}$ denote the sum over all m-tuples $(j_1, \ldots, j_m)$ that sum to $n-(m-1)$ and $\sum_{\mathbf{k}}$ the sum over all $(m+1)$-tuples $(k_1, \ldots, k_{m+1})$ of nonnegative integers that sum to k. Then, letting $\oplus$ denote sums of

mutually independent random variables, for $n \geq m-1$ we have

$$\mu_n(k) = \mathbf{E}\, X_n^k = \mathbf{E}\,\mathbf{E}\,(X_n^k \mid |L_1|, \ldots, |L_m|) = \frac{1}{\binom{n}{m-1}} \sum_{\mathbf{j}} \mathbf{E}\,(X_{j_1} \oplus \cdots \oplus X_{j_m} + t_n)^k$$

$$= \frac{1}{\binom{n}{m-1}} \sum_{\mathbf{j}} \sum_{\mathbf{k}} \binom{k}{k_1, \ldots, k_m, k_{m+1}} \mu_{j_1}(k_1) \cdots \mu_{j_m}(k_m) t_n^{k_{m+1}}.$$

We can rewrite this equation as

$$\mu_n(k) = \frac{m}{\binom{n}{m-1}} \sum_{j=0}^{n-(m-1)} \binom{n-1-j}{m-2} \mu_j(k) + r_n(k), \tag{2.1}$$

where

$$r_n(k) := \sum_{\mathbf{k}}{}^{*} \binom{k}{k_1, \ldots, k_m, k_{m+1}} t_n^{k_{m+1}} \frac{1}{\binom{n}{m-1}} \sum_{\mathbf{j}} \mu_{j_1}(k_1) \cdots \mu_{j_m}(k_m), \tag{2.2}$$

with $\sum_{\mathbf{k}}^{*}$ denoting the same sum as $\sum_{\mathbf{k}}$ with the additional restriction that $k_i < k$ for $i = 1, \ldots, m$. We have thus established that the moments $\mu_n(k)$ each satisfy the same basic recurrence in n, differing as k varies only in the non-homogeneous term $r_n(k)$. Observe that $r_n(1) = t_n$, the toll function. We record this important fact as

Proposition 2.1. *Under the random permutation model, all moments of an additive functional satisfy the basic recurrence*

$$a_n = b_n + \frac{m}{\binom{n}{m-1}} \sum_{j=0}^{n-(m-1)} \binom{n-1-j}{m-2} a_j, \qquad n \geq m-1, \tag{2.3}$$

with specified initial conditions (say) $a_j := b_j$, $0 \leq j \leq m-2$. *[Recall the statement following Definition 1.1 about the initial conditions for the recurrence* (1.1)*.]*

To be more specific, equation (2.3) is satisfied by $a_n = \mu_n(k) = \mathbf{E}\, X_n^k$ and $b_n = r_n(k)$, where $r_n(k)$ is defined in terms of lower-order moments of smaller trees at (2.2). We proceed to study the recurrence relation (2.3) for general input (b_n) and corresponding output (a_n).

2.2 Transfer theorems

In order to analyze the recurrence relation (2.3) we introduce generating functions

$$A(z) := \sum_{n=0}^{\infty} a_n z^n \quad \text{and} \quad B(z) := \sum_{n=0}^{\infty} b_n z^n.$$

Furthermore, let $x^{\overline{r}} := \prod_{k=0}^{r-1}(x+k)$ denote the rth rising factorial power of x and $x^{\underline{r}} := \prod_{k=0}^{r-1}(x-k)$ the rth falling factorial power of x. Multiplying (2.3) by $n^{\underline{m-1}} z^{n-(m-1)}$

and summing over $n \geq m-1$ we get, after some (routine) calculation, the differential equation

$$A^{(m-1)}(z) = B^{(m-1)}(z) + m!(1-z)^{-(m-1)}A(z). \tag{2.4}$$

Equations of the form (2.4) are members of a class known as *Euler differential equations.* Using the method of variation of parameters and combinatorial identities one can obtain the general solution to this equation. See Section 3 for a proof.

Theorem 2.2 (Exact Transfer Theorem (ETT)). *Let A and B denote the respective ordinary generating functions of the sequences (a_n) and (b_n) in the recurrence* (2.3)*. Let*

$$\hat{B}(z) := B(z) - \sum_{j=0}^{m-2} b_j z^j = \sum_{n=m-1}^{\infty} b_n z^n. \tag{2.5}$$

Then

$$\begin{aligned} A(z) &= \sum_{j=1}^{m-1} c_j (1-z)^{-\lambda_j} + \sum_{j=1}^{m-1} \frac{(1-z)^{-\lambda_j}}{\psi'(\lambda_j)} \int_0^z B^{(m-1)}(\zeta)(1-\zeta)^{\lambda_j+m-2}\, d\zeta && (2.6)\\ &= \sum_{j=1}^{m-1} c_j (1-z)^{-\lambda_j} + \hat{B}(z) + m! \sum_{j=1}^{m-1} \frac{(1-z)^{-\lambda_j}}{\psi'(\lambda_j)} \int_0^z \hat{B}(\zeta)(1-\zeta)^{\lambda_j-1}\, d\zeta, && (2.7) \end{aligned}$$

where ψ is the indicial polynomial

$$\psi(\lambda) := \lambda^{\overline{m-1}} - m! = \lambda(\lambda+1)\cdots(\lambda+m-2) - m! \tag{2.8}$$

with roots $2 =: \lambda_1, \lambda_2, \ldots, \lambda_{m-1}$ in nonincreasing order of real parts. In (2.6)*, the coefficients $c_1, c_2, \ldots, c_{m-1}$ can be written explicitly in terms of the initial conditions $b_0, \ldots, b_{m-2}$ as*

$$c_j = \frac{m!}{\psi'(\lambda_j)} \sum_{k=0}^{m-2} b_k \frac{k!}{\lambda_j^{\overline{k+1}}}, \qquad j = 1, \ldots, m-1. \tag{2.9}$$

In particular,

$$c_1 = \frac{1}{H_m - 1} \sum_{j=0}^{m-2} \frac{b_j}{(j+1)(j+2)}. \tag{2.10}$$

The indicial polynomial (2.8) is well studied; see [18, 17, 5] and Appendix B, and also the related [19]. We will exploit the expression (2.6) for $A(z)$ to relate the asymptotic properties of the sequence (b_n) to those of (a_n) and then use *transfer theorems* to derive limiting distributions of additive functionals.

Remark 2.3. For computations, equation (2.6) might be easier to use when it is no bother to compute derivatives of B; otherwise, (2.7) is easier. Equation (2.7) will be used in establishing part (a) of the Asymptotic Transfer Theorem 2.4; the proof of part (b) will use (2.6).

It is quite easy to transfer asymptotics for B to asymptotics for A using the ETT. We give three examples important for applications to moments of functionals in the next theorem, proved independently by Félix Chern and Hsien-Kuei Hwang in [5] using a quite different approach. The series convergence required in (2.11) need not be absolute.

Theorem 2.4 (Asymptotic Transfer Theorem (ATT)).

(a) *If*

$$b_n = o(n) \qquad \text{and} \qquad \sum_{n=0}^{\infty} \frac{b_n}{(n+1)(n+2)} \text{ converges,} \tag{2.11}$$

then

$$a_n = \frac{K_1}{H_m - 1} n + o(n), \qquad \text{where} \qquad K_1 := \sum_{j=0}^{\infty} \frac{b_j}{(j+1)(j+2)}. \tag{2.12}$$

(b) *If $b_n \equiv K_2(n+1) + h_n$ where (h_n) satisfies (2.11) [with (b_n) replaced by (h_n)], then*

$$a_n = \frac{K_2}{H_m - 1} n H_n + \frac{K_3}{H_m - 1} n + o(n), \tag{2.13}$$

where

$$K_3 := \sum_{j=0}^{\infty} \frac{h_j}{(j+1)(j+2)} + K_2 \left[\frac{H_m - 1}{2} - 1 + \frac{H_m^{(2)} - 1}{2(H_m - 1)} \right]. \tag{2.14}$$

(c) *If $b_n = K_4 n^v + o(n^v)$ with $v > 1$, then*

$$a_n = \frac{K_4}{1 - \frac{m!\Gamma(v+1)}{\Gamma(v+m)}} n^v + o(n^v). \tag{2.15}$$

Of course part (a) is equally valid with $\sum_{n=0}^{\infty} \frac{b_n}{n^2}$ replacing the series in (2.11).

Theorem 2.4 will be proved in Section 4. It is easy to see that the ratio $\frac{m!\Gamma(v+1)}{\Gamma(v+m)}$ appearing in (2.15) has modulus strictly less than unity; in particular, the expression given is well defined. Refined and additional asymptotic transfers will be discussed in Section 5.

2.3 Limiting distributions

Applications to moments and limiting distributions are discussed for "small" toll sequences in Section 5—see Theorems 5.1 and 5.4; "moderate" and "large" toll sequences are discussed in Section 6—see Theorems 6.2 and 6.5. We state here a summary theorem that can be deduced easily from these four theorems. [For the definition of $m_0(\beta)$ in case (ii) of the theorem, see Remark 6.3. Concerning the values of g_1 and g_2 in cases (ii) and (iii), consult (6.5), (6.6) (and (6.2)).]

Theorem 2.5 (Limit theorem for additive functionals). *Let $X_n = f(T_n)$ be the additive functional on random m-ary search trees corresponding to a toll sequence (t_n),*

with T_n having the distribution on trees induced by a uniformly distributed permutation on $[n]$. When (t_n) satisfies one of the three conditions (i)–(iii) *below, then*

$$\frac{X_n - \mathbf{E}\,X_n}{\sqrt{\mathbf{Var} X_n}} \xrightarrow{\mathcal{L}} W,$$

with convergence of all moments. Here, with L denoting a slowly varying function,

(i) *If $2 \le m \le 26$ and*

   (a) $t_n = o(\sqrt{n})$ *and* $\sum^{\infty} n^{-1} \max_{n^\delta \le k \le n} \frac{t_k^2}{k^2} < \infty$ *for some* $0 < \delta < 1$ *or*

   (b) $t_n \sim \sqrt{n} L(n)$,

   then W has the standard normal distribution.

(ii) *If $t_n \sim n^\beta L(n)$, with $1/2 < \beta < 1$ and $2 \le m \le m_0(\beta)$, then $W = g_2^{-1/2} Y$ where $\mathcal{L}(Y)$ is the unique distribution with finite second moment satisfying* (6.1)*. Here*

$$g_2 = \frac{(m-1)!}{1 - \frac{m!\Gamma(2\beta+1)}{\Gamma(2\beta+m)}} \left[\frac{1}{(m-1)!} + 2mg_1 \frac{\Gamma(\beta+1)}{\Gamma(\beta+m)} + m(m-1)g_1^2 \frac{\Gamma^2(\beta+1)}{\Gamma(2\beta+m)} \right] > 0, \tag{2.16}$$

   with

$$g_1 = \left(1 - \frac{m!\Gamma(\beta+1)}{\Gamma(\beta+m)}\right)^{-1}. \tag{2.17}$$

(iii) *If $t_n \sim n^\beta L(n)$ with $\beta > 1$, then $W = g_2^{-1/2} Y$ where $\mathcal{L}(Y)$ is the unique distribution with finite second moment satisfying* (6.1)*, where g_2 is again defined at* (2.16)–(2.17)*.*

Remark 2.6. In case (iii), it is easy to check that $g_1 > 0$ and $g_2 > 0$. In case (ii) one can verify easily that $g_1 < 0$. In this case to see that $g_2 > 0$, note that no constant random variable satisfies (6.1).

3 Proof of the ETT

In this section we prove the ETT, which is Theorem 2.2.

Proof. For the proof of (2.6), see Appendix A, in particular Corollary A.2 and Proposition A.3. To begin the proof of (2.7), note that B can be replaced by $\hat{B}$ in (2.6). We then use repeated integration by parts and invoke Identity B.10. Denoting

$$\hat{A} := A(z) - \sum_{j=1}^{m-1} c_j (1-z)^{-\lambda_j},$$

after $m-2$ integrations by parts we find

$$\hat{A}(z) = \sum_{j=1}^{m-1} \frac{(1-z)^{-\lambda_j}}{\psi'(\lambda_j)} (\lambda_j + m - 2) \cdots (\lambda_j + 1) \int_0^z \hat{B}'(\zeta)(1-\zeta)^{\lambda_j}\, d\zeta.$$

But
$$(\lambda_j + m - 2) \cdots (\lambda_j + 1) = \frac{\lambda_j^{\overline{m-1}}}{\lambda_j} = \frac{\psi(\lambda_j) + m!}{\lambda_j} = \frac{m!}{\lambda_j},$$
so
$$\hat{A}(z) = m! \sum_{j=1}^{m-1} \frac{(1-z)^{-\lambda_j}}{\lambda_j \psi'(\lambda_j)} \int_0^z \hat{B}'(\zeta)(1-\zeta)^{\lambda_j}\, d\zeta.$$
We obtain (2.7) by performing one more integration by parts and utilizing Identity B.9 with $\lambda = 0$. □

4 Proof of the ATT

In this section we prove the ATT, which is Theorem 2.4. The following elementary result, which is found in the first line of the proof of Lemma 6 in [5], is key to the analysis of (2.7). For completeness, we include a proof.

Lemma 4.1. *Let $Y(z) = \sum_{n=0}^{\infty} y_n z^n$ with $y_0 = 0$. For any $\lambda \in \mathbf{C}$,*
$$[z^n] \left((1-z)^{-\lambda} \int_0^z (1-\zeta)^{\lambda-1} Y(\zeta)\, d\zeta \right) = \sum_{k=0}^{n-1} \frac{y_k}{k+1} \prod_{j=k+2}^{n} \left(1 + \frac{\lambda - 1}{j} \right), \qquad n \geq 0. \tag{4.1}$$
The product in (4.1) may be written (when $\lambda \in \mathbf{C} \setminus \{0, -1, -2, \ldots\}$) as
$$\frac{\Gamma(\lambda+n)\Gamma(2+k)}{\Gamma(1+n)\Gamma(\lambda+k+1)},$$
which by Stirling's formula equals
$$\frac{n^{\lambda-1}\left[1 + O(n^{-1})\right]}{(k+1)^{\lambda-1}\left[1 + O((k+1)^{-1})\right]} \tag{4.2}$$
for $n \geq 1$ and $k \geq 1$. [The product in (4.1) equals (4.2) even if $\lambda \in \{0, -1, -2, \ldots\}$.] Also, of special interest is the case $\lambda = 2$, in which case (4.1) reduces to
$$[z^n] \left((1-z)^{-2} \int_0^z (1-\zeta) Y(\zeta)\, d\zeta \right) = (n+1) \sum_{k=0}^{n-1} \frac{y_k}{(k+1)(k+2)}, \qquad n \geq 0. \tag{4.3}$$

Proof. The function $W(z) := (1-z)^{-\lambda} \int_0^z (1-\zeta)^{\lambda-1} Y(\zeta)\, d\zeta$ is the unique solution with $W(0) = 0$ to the differential equation
$$W'(z) = \lambda(1-z)^{-1} W(z) + (1-z)^{-1} Y(z);$$
that is, $w_n := [z^n]W(z)$, $n \geq 0$, satisfies $w_0 = 0$ and
$$w_n = \frac{\lambda}{n} \sum_{k=0}^{n-1} w_k + \frac{1}{n} \sum_{k=0}^{n-1} y_k, \qquad n \geq 1. \tag{4.4}$$
But the recurrence (4.4) can be easily solved to yield (4.1): compute the difference $n w_n - (n-1) w_{n-1}$ and iterate. □

For part (a) of the ATT we use the following estimates from [5]; these follow readily from (4.1)–(4.3). [Part (a) is their Lemma 6; part (b) is used tacitly in the proof of their Proposition 7.]

Lemma 4.2.

(a) *If* $\mathrm{Re}(\lambda) < 2$ *and* $Y(z) = \sum_{n=0}^{\infty} y_n z^n$ *satisfies* $y_0 = 0$ *and* $y_n = o(n)$, *then*

$$[z^n]\left((1-z)^{-\lambda}\int_0^z Y(\zeta)(1-\zeta)^{\lambda-1}\,d\zeta\right) = o(n).$$

(b) *With* $\hat{B}$ *defined at* (2.5), *if* (2.11) *holds, then*

$$[z^n]\left((1-z)^{-2}\int_0^z \hat{B}(\zeta)(1-\zeta)\,d\zeta\right) = n\sum_{j=m-1}^{\infty}\frac{b_j}{(j+1)(j+2)} + o(n). \qquad \square$$

For part (c) of the ATT, we will need the following simple comparison lemma.

Lemma 4.3. *If* (b_n) *and* (b'_n) *are two input sequences such that*

$$|b_n| \le b'_n \quad \text{for all } n \ge 0,$$

then the corresponding output sequences (a_n) *and* (a'_n) *in* (2.3) *(with the initial conditions stated there) satisfy*

$$|a_n| \le a'_n \quad \text{for all } n \ge 0.$$

Proof. This follows immediately by induction. $\square$

Proof of Theorem 2.4. (a) From (2.7), assumption (2.11), Lemma 4.2, (2.10), and $\psi'(2) = m!(H_m - 1)$, the result is immediate.

(b) Suppose first that $b_n \equiv n+1$. Then $B(z) \equiv (1-z)^{-2}$, so

$$B^{(m-1)}(z) \equiv m!(1-z)^{-(m+1)}.$$

Plugging this into (2.6) we find

$$a_n = (n+1)\left[c_1 + m!\sum_{j=2}^{m-1}\frac{1}{(2-\lambda_j)\psi'(\lambda_j)}\right] + \frac{m!}{\psi'(2)}[z^n]\left[(1-z)^{-2}\log\left((1-z)^{-1}\right)\right] + o(n).$$

Now we use Identities B.11 and B.12. Also, since $b_n \equiv n+1$, we have from (2.10) in this case that $c_1 = 1$. Therefore

$$\begin{aligned} a_n &= (n+1)\left(1 + \frac{1}{2}\left[\frac{H_m^{(2)} - 1}{(H_m-1)^2} - 1\right]\right) + \frac{1}{H_m - 1}[(n+1)H_n - n] + o(n) \\ &= \frac{1}{H_m-1}nH_n + \left[\frac{1}{2} - \frac{1}{H_m-1} + \frac{H_m^{(2)}-1}{2(H_m-1)^2}\right]n + o(n). \end{aligned}$$

This completes the proof of (b) for our special case, and the general case follows from this and part (a) using the superposition principle.

(c) Suppose first that $b_n \equiv (v+1)^{\overline{n}}/n! \sim n^v/\Gamma(v+1)$, so that $B(z) \equiv (1-z)^{-(v+1)}$ and $B^{(m-1)}(z) \equiv (v+1)^{\overline{m-1}}(1-z)^{-(v+m)}$. Plugging this into (2.6) and utilizing Identity B.9 with $\lambda = v+1$ and the calculation

$$\frac{(v+1)^{\overline{m-1}}}{(v+1)^{\overline{m-1}} - m!} = \left[1 - \frac{m!\Gamma(v+1)}{\Gamma(v+m)}\right]^{-1},$$

we find

$$A(z) = \left[1 - \frac{m!\Gamma(v+1)}{\Gamma(v+m)}\right]^{-1} (1-z)^{-(v+1)} + O(|1-z|^{-2}).$$

By singularity analysis for large functions [12], this completes the proof of (c) for our special case.

To complete the proof in the general case, we need only show that if $b_n = o(n^v)$ for $v > 1$, then $a_n = o(n^v)$. Indeed, fix $\varepsilon > 0$; then there exists a sequence (b'_n) such that $|b_n| \le b'_n$ for all n and

$$b'_n = \varepsilon(v+1)^{\overline{n}}/n! \quad \text{for all large } n.$$

The toll sequence is but a slight modification of our special-case toll sequence, and we see that

$$a'_n = \varepsilon' n^v + o(n^v), \quad \text{where} \quad \varepsilon' := \frac{\varepsilon}{\Gamma(v+1)}\left[1 - \frac{m!\Gamma(v+1)}{\Gamma(v+m)}\right]^{-1}.$$

Now Lemma 4.3 implies that

$$\limsup_n |a_n| n^{-v} \le \varepsilon';$$

since ε (and hence ε') can be made arbitrarily small, this completes the proof. □

The conditions (2.11) on (b_n) are not only sufficient but also necessary for asymptotic linearity of a_n. Indeed, here is a converse:

Proposition 4.4. *If $a_n = Kn + o(n)$ for some constant K, then* (2.11) *holds.*

Proof. Chern and Hwang [5] provide the simple proof that $b_n = o(n)$. Moreover, then, from (2.7), (2.10), and (4.3) with Y taken to be $\hat{B}$, we find that

$$\sum_{j=0}^{n-1} \frac{b_j}{(j+1)(j+2)} = K(H_m - 1) + o(1),$$

i.e., that the series $\sum_{n=0}^{\infty} \frac{b_n}{(n+1)(n+2)}$ converges [to $K(H_m - 1)$]. □

The following additional asymptotic transfer results are established by calculations similar to those in the proof of the ATT. We leave detailed proofs as exercises for the reader.

Theorem 4.5 (more asymptotic transfers). *Consider the initial value problem* (2.3)*.*

(a) *If* $2 \le m \le 26$ *and* $b_n = o(\sqrt{n})$*, then we can refine* (2.12) *to*

$$a_n = \frac{K_1}{H_m - 1} n + o\left(\sqrt{n}\right). \tag{4.5}$$

(b) *If* $\operatorname{Re} \lambda_2 =: \alpha < 1 + \beta$ *and* $b_n \sim n^\beta L(n)$*, where* $1/2 < \beta < 1$*, with* L *slowly varying, then we can refine* (2.12) *to*

$$a_n = \frac{K_1}{H_m - 1} n - \frac{(1+\beta)^{\overline{m-1}}}{m! - (1+\beta)^{\overline{m-1}}} n^\beta L(n) + o\left(n^\beta L(n)\right). \tag{4.6}$$

(c) *If* $b_n \sim nL(n)$ *with* L *slowly varying, then, with* K_1 *defined at* (2.12)*,*

$$a_n \sim \begin{cases} \frac{K_1}{H_m-1} n, & \text{if } \sum^\infty \frac{L(k)}{k} < \infty, \\ \frac{1}{H_m-1} n \sum_{k \le n} \frac{L(k)}{k}, & \text{if } \sum^\infty \frac{L(k)}{k} = \infty. \end{cases} \tag{4.7}$$

(d) *Part* (c) *of the* ATT *can be extended as follows. If* $b_n = K_4 n^v L(n) + o\left(n^v L(n)\right)$ *with* $v > 1$ *and* L *slowly varying, then*

$$a_n = \frac{K_4}{1 - \frac{m!\Gamma(v+1)}{\Gamma(v+m)}} n^v L(n) + o\left(n^v L(n)\right). \tag{4.8}$$

Proof hints. Whenever the conditions (2.11) are met we have by the ETT and (4.3)

$$\begin{aligned} a_n - \frac{K_1}{H_m - 1}(n+1) \quad &= \quad O(n^{\alpha - 1}) + b_n - \frac{1}{H_m - 1}(n+1) \sum_{k=n}^\infty \frac{\hat{b}_k}{(k+1)(k+2)} \\ &\quad + m! \sum_{j=2}^{m-1} \frac{1}{\psi'(\lambda_j)} [z^n] \left((1-z)^{-\lambda_j} \int_0^z \hat{B}(\zeta)(1-\zeta)^{\lambda_j - 1}\, d\zeta \right), \end{aligned} \tag{4.9}$$

where α is strictly smaller than $1 + \beta$ by assumption. (When $\beta = 1/2$ we know that $\alpha < 3/2$ when $m \le 26$.) Simple estimates, including the use of (4.2), give cases (a) and (b); for (b), the coefficient of $n^\beta L(n)$ in (4.6) indeed is, using Identities B.9 and B.11,

$$1 - \frac{1}{(1-\beta)(H_m - 1)} + m! \sum_{j=2}^{m-1} \frac{1}{((1+\beta) - \lambda_j)\psi'(\lambda_j)} = -\frac{(1+\beta)^{\overline{m-1}}}{m! - (1+\beta)^{\overline{m-1}}}.$$

In case (c), from the ETT result (2.7) and simple estimates we find

$$a_n = (n+1) \frac{1}{H_m - 1} \sum_{k=0}^{n-1} \frac{b_k}{(k+1)(k+2)} + O(nL(n)).$$

To finish, we use the regular variation fact (quoted by Hwang and Neininger [14] at their equation (7) from Proposition 1.5.9a in [2]) that

$$L(n) = o\left(\sum_{k \le n} \frac{L(k)}{k}\right). \tag{4.10}$$

In case (d), again from (2.7) and simple estimates we find

$$a_n = O(n) + b_n + K_4 n^v L(n) m! \sum_{j=1}^{m-1} \frac{1}{\psi'(\lambda_j)(v+1-\lambda_j)} + o\left(n^v L(n)\right).$$

The proof is completed by using Identity B.9. □

5 Asymptotic normality for small toll functions

In this section we establish asymptotic normality for "small" toll functions.

5.1 Central limit theorem statements

As did Hwang and Neininger in [14], in our Theorems 5.1 and 5.4 we treat two overlapping cases. Throughout, we write $\sum_{\mathbf{j}}$ as shorthand for the sum over m-tuples $(j_1, \dots, j_m)$ of nonnegative integers summing to $n - (m-1)$.

Theorem 5.1 (CLT I for small toll functions). *If $2 \le m \le 26$ and the toll sequence (t_n) satisfies*

$$\text{(a) } t_n = o\left(\sqrt{n}\right) \qquad \text{and} \qquad \text{(b) } \sum^{\infty} n^{-1} \max_{n^\delta \le k \le n} \frac{t_k^2}{k} < \infty \text{ for some } 0 < \delta < 1, \tag{5.1}$$

then the mean μ_n and variance σ_n^2 of the corresponding additive functional X_n on m-ary search trees with the random permutation model satisfy, respectively,

$$\mu_n = \frac{K_1}{H_m - 1} n + o\left(\sqrt{n}\right) =: \mu n + o\left(\sqrt{n}\right), \tag{5.2}$$

with K_1 defined at (2.12), and

$$\sigma_n^2 = \sigma^2 n + o(n), \qquad \text{where} \qquad \sigma^2 := \frac{1}{H_m - 1} \sum_{j=0}^{\infty} \frac{r_j}{(j+1)(j+2)}, \tag{5.3}$$

with the sequence (r_n) defined by $r_j := 0$ for $0 \le j \le m-2$ and

$$r_n := \frac{1}{\binom{n}{m-1}} \sum_{\mathbf{j}} \left[t_n + \mu_{j_1} + \cdots + \mu_{j_m} - \mu_n\right]^2, \qquad n \ge m-1. \tag{5.4}$$

Moreover,

$$\frac{X_n - \mu n}{\sqrt{n}} \text{ is asymptotically } \mathrm{N}(0, \sigma^2),$$

and there is convergence of moments of every order.

Remark 5.2. One can check (for *any* $2 \le m < \infty$) that the variance σ_n^2 vanishes for all $n \ge m-1$ if and only if the toll sequence is chosen as

$$t_n = t \min\{m-1, n\}, \qquad n \ge 0$$

for some constant $t \in \mathbf{R}$. In that case without loss of generality $t = 1$ and then $X_n \equiv n$ is just the number of keys and we have exact (though degenerate) normality. So we shall assume in the proof of Theorem 5.1 that $\sigma^2 > 0$.

Remark 5.3. (a) Condition (5.1)(b) trivially implies

$$\sum^{\infty} \frac{t_n^2}{n^2} < \infty, \tag{5.5}$$

which in turn implies that (2.11) holds with absolute convergence; indeed, since the nonnegative numbers $[(n+1)(n+2)]^{-1}$, $n \ge 0$, sum to unity, we have

$$\left[\sum_{n=0}^{\infty} \frac{|t_n|}{(n+1)(n+2)}\right]^2 \le \sum_{n=0}^{\infty} \frac{t_n^2}{(n+1)(n+2)} < \infty. \tag{5.6}$$

(b) If

$$|t_n| = O\left(\tilde{t}_n\right) \quad \text{with} \quad 0 \le \frac{\tilde{t}_n}{\sqrt{n}} \downarrow \quad \text{and} \quad \sum^{\infty} \frac{\tilde{t}_n^2}{n^2} < \infty, \tag{5.7}$$

then we claim that (5.1) holds, and then as a corollary

$$\frac{t_n}{\sqrt{n}} \downarrow 0 \quad \text{and} \quad \sum^{\infty} \frac{t_n^2}{n^2} < \infty$$

implies (5.1). To see the claim, first observe that the condition (5.7) certainly implies (5.1)(a); moreover, we observe that the series (say, over $2 \le n < \infty$) in (5.1)(b) is bounded by a constant times

$$\sum_{n=2}^{\infty} n^{-1} \max_{n^\delta \le k \le n} \frac{\tilde{t}_k^2}{k} = \sum_{n=2}^{\infty} n^{-1} \frac{\tilde{t}^2_{\lceil n^\delta \rceil}}{\lceil n^\delta \rceil} = \sum_{k=2}^{\infty} \frac{\tilde{t}_k^2}{k} \sum_{(k-1)^{1/\delta} < n \le k^{1/\delta}} n^{-1} = O\left(\sum_{k=2}^{\infty} \frac{\tilde{t}_k^2}{k^2}\right) < \infty.$$

(c) One can check that, when $m = 2$, the proof we will give of CLT I requires only (5.1)(a) and (5.5). In that case we obtain a strengthening of "Case S1" of Theorem 2 in [14] (for deterministic toll sequences); they required $t_n = O\left(\sqrt{n}/(\log n)^{(1/2)+\varepsilon}\right)$ for some $\varepsilon > 0$.

Theorem 5.4 (CLT II for small toll functions). *If* $2 \le m \le 26$ *and the toll sequence* (t_n) *satisfies*

$$t_n \sim \sqrt{n} L(n) \tag{5.8}$$

with L *slowly varying, then the mean* μ_n *of the corresponding additive functional* X_n *on* m*-ary search trees with the random permutation model satisfies*

$$\mu_n = \frac{K_1}{H_m - 1} n - \frac{(3/2)^{\overline{m-1}}}{m! - (3/2)^{\overline{m-1}}} \sqrt{n} L(n) + o\left(\sqrt{n} L(n)\right). \tag{5.9}$$

with K_1 defined at (2.12). If $\sum^{\infty} \frac{L^2(k)}{k} < \infty$, then the variance σ_n^2 satisfies (5.3)–(5.4) *and we define*

$$s^2(n) := \sigma^2 n.$$

If $\sum^{\infty} \frac{L^2(k)}{k} = \infty$, then

$$\sigma_n^2 \sim s^2(n) := \sigma^2 n \sum_{k \le n} \frac{L^2(k)}{k} \tag{5.10}$$

where in this case we define

$$\sigma^2 := \frac{\left((\tfrac{3}{2})^{\overline{m-1}} \right)^2 \left[\tfrac{\pi}{4}(m-1) + 1 \right] - (m!)^2}{(H_m - 1) \left[m! - (\tfrac{3}{2})^{\overline{m-1}} \right]^2}. \tag{5.11}$$

Moreover, in either case

$$\frac{X_n - \mu n}{s(n)} \text{ is asymptotically standard normal,}$$

and there is convergence of moments of every order.

Remark 5.5. When $m = 2$ the constant σ^2 in (5.11) equals $\frac{9}{2}\pi - 14$, and Theorem 5.4 reduces to "Case S2" of Theorem 2 in [14] (for deterministic toll sequences): see especially their displays (15) and (17), with $\tau_2 = 1$.

5.2 Central limit theorem proofs

Proof of CLT I (Theorem 5.1). We use the method of moments together with asymptotic transfer results.

Given the toll sequence (t_n) defining the sequence (X_n) of random functionals of interest, the means (μ_n) satisfy the basic recurrence relation (2.3) with (b_n) replaced by (t_n). Thus (5.2) simply repeats the asymptotic transfer result (4.5).

According to the law of total variance, the sequence (σ_n^2) of variances also satisfies the recurrence (2.3), but with (b_n) replaced by (r_n). According to Lemma 5.6 to follow, the sequence (r_n) satisfies the conditions (2.11). [Note: When $m = 2$, only (5.1)(a) and (5.5), not (5.1)(b), are needed in the proof of Lemma 5.6.] Then (5.3) is immediate from part (a) of the ATT.

Let $\tilde{X}_n := X_n - \mu(n+1)$ for $n \ge 0$. We will complete the proof of CLT I by showing by induction on k that

$$\tilde{\mu}_n(k) := \mathbf{E}\, \tilde{X}_n^k, \qquad k \ge 1 \qquad [\text{with } \tilde{\mu}_n(0) := 1] \tag{5.12}$$

satisfies

$$\tilde{\mu}_n(2k) \sim \frac{(2k)!}{2^k k!} \sigma^{2k} n^k, \qquad k \ge 1 \tag{5.13}$$

and

$$\tilde{\mu}_n(2k-1) = o\left(n^{k-(1/2)}\right), \qquad k \ge 1. \tag{5.14}$$

Observe that (5.2)–(5.3) imply that (5.13) and (5.14) both hold for $k = 1$.

The key to the induction step for both (5.13) and (5.14) is to apply the law of total expectation to (5.12), by conditioning on the subtree sizes $|L_1|, \ldots, |L_m|$ (recall the notation in Section 1). In a manner similar to (2.1), we have

$$\tilde{\mu}_n(k) = \frac{m}{\binom{n}{m-1}} \sum_{j=0}^{n-(m-1)} \binom{n-1-j}{m-2} \tilde{\mu}_j(k) + r_n(k), \tag{5.15}$$

where

$$r_n(k) := \sum_{\mathbf{k}}{}^{*} \binom{k}{k_1, \ldots, k_m, k_{m+1}} t_n^{k_{m+1}} \times \frac{1}{\binom{n}{m-1}} \sum_{\mathbf{j}} \tilde{\mu}_{j_1}(k_1) \cdots \tilde{\mu}_{j_m}(k_m) \tag{5.16}$$

with $\sum_{\mathbf{k}}^{*}$ here denoting the same sum as $\sum_{\mathbf{k}}$ with the additional restriction that $k_i < k$ for $i = 1, \ldots, m$. Observe that (5.15) is again of the basic form (2.3). We will apply the ATT after evaluating $r_n(k)$ asymptotically.

We will treat the induction step in detail only for (5.13), the case (5.14) being similar and somewhat easier. Let $\sum_{\mathbf{k}}^{**}$ denote the sum over m-tuples $(k_1, \ldots, k_m)$ of nonnegative integers, each $< k$, summing to k (i.e., the same sum as $\sum_{\mathbf{k}}^{*}$ with the additional restriction that $k_{m+1} = 0$). For $k \geq 2$ we clearly have, by induction [recall (5.13)–(5.14)],

$$\begin{aligned}
r_n(2k) &= o(n^k) + \sum_{\mathbf{k}}{}^{**} \binom{2k}{2k_1, \ldots, 2k_m} \frac{1}{\binom{n}{m-1}} \sum_{\mathbf{j}} \tilde{\mu}_{j_1}(2k_1) \cdots \tilde{\mu}_{j_m}(2k_m) \\
&= o(n^k) + \sum_{\mathbf{k}}{}^{**} \binom{2k}{2k_1, \ldots, 2k_m} \frac{1}{\binom{n}{m-1}} \sum_{\mathbf{j}} \frac{(2k_1)!}{2^{k_1} k_1!} \sigma^{2k_1} j_1^{k_1} \cdots \frac{(2k_m)!}{2^{k_m} k_m!} \sigma^{2k_m} j_m^{k_m} \\
&= o(n^k) + \frac{(2k)!}{2^k k!} \sigma^{2k} n^k \sum_{\mathbf{k}}{}^{**} \binom{k}{k_1, \ldots, k_m} \frac{1}{\binom{n}{m-1}} \sum_{\mathbf{j}} \left(\frac{j_1}{n}\right)^{k_1} \cdots \left(\frac{j_m}{n}\right)^{k_m}.
\end{aligned}$$

But

$$\begin{aligned}
&\frac{1}{\binom{n}{m-1}} \sum_{\mathbf{j}} \left(\frac{j_1}{n}\right)^{k_1} \cdots \left(\frac{j_m}{n}\right)^{k_m} \\
&\quad \to (m-1)! \int x_1^{k_1} \cdots x_{m-1}^{k_{m-1}} (1 - x_1 - \cdots x_{m-1})^{k_m} \, dx_1 \cdots dx_{m-1} \\
&\quad = (m-1)! \frac{\Gamma(k_1+1) \cdots \Gamma(k_m+1)}{\Gamma(k+m)} = \frac{1}{\binom{k}{k_1, \ldots, k_m} \binom{k+m-1}{m-1}},
\end{aligned}$$

where the above integral is over $(x_1, \ldots, x_{m-1}) \in [0,1]^{m-1}$ with sum not exceeding unity. Since the number of terms in $\sum_{\mathbf{k}}^{**}$ is $\binom{k+m-1}{m-1} - m$, we therefore have

$$\begin{aligned}
r_n(2k) &= \frac{(2k)!}{2^k k!} \sigma^{2k} n^k \frac{\binom{k+m-1}{m-1} - m}{\binom{k+m-1}{m-1}} + o(n^k) \\
&= \frac{(2k)!}{2^k k!} \sigma^{2k} n^k \left[1 - \frac{m! \Gamma(k+1)}{\Gamma(k+m)}\right] + o(n^k), \qquad k \geq 2.
\end{aligned}$$

Similarly,

$$r_n(2k-1) = o\left(n^{k-(1/2)}\right), \qquad k \geq 2.$$

Now part (c) of the ATT implies (5.13) and (5.14). □

The following lemma lies at the heart of the proof of Theorem 5.1.

Lemma 5.6. *In the context of CLT I, the sequence* (r_n) *defined at* (5.4) *satisfies the conditions* (2.11)*.*

Proof. Clearly

$$r_n = \frac{1}{\binom{n}{m-1}} \sum_{\mathbf{j}} \left[t_n + \tilde{\mu}_{j_1} + \cdots + \tilde{\mu}_{j_m} - \tilde{\mu}_n\right]^2, \qquad n \geq m-1, \tag{5.17}$$

with

$$\tilde{\mu}_n := \mu_n - \mu(n+1), \tag{5.18}$$

which is $o(\sqrt{n})$ by (5.2). Recall the inequality

$$\left[\sum_{i=1}^{k} \xi_i\right]^2 \leq k \sum_{i=1}^{k} \xi_i^2 \tag{5.19}$$

for real numbers $\xi_1, \ldots, \xi_k$. Applying this to (5.17),

$$\frac{r_n}{m+2} \leq t_n^2 + \tilde{\mu}_n^2 + \frac{m}{\binom{n}{m-1}} \sum_{j=0}^{n-(m-1)} \binom{n-1-j}{m-2} \tilde{\mu}_j^2, \tag{5.20}$$

from which (2.11)(a) for (r_n) is evident.

To establish the summability of r_n/n^2 we need only establish that of $\tilde{\mu}_n^2/n^2$. Indeed we can then use (5.20) again, together with (5.5) and the estimate

$$\begin{aligned}
&\sum_{n=m-1}^{\infty} n^{-2} \frac{m}{\binom{n}{m-1}} \sum_{j=0}^{n-(m-1)} \binom{n-1-j}{m-2} \tilde{\mu}_j^2 \\
&= m(m-1) \sum_{j=0}^{\infty} \tilde{\mu}_j^2 \sum_{n=j+m-1}^{\infty} \frac{(n-1-j)_{m-2}}{n^2 (n)_{m-1}} \\
&\leq m(m-1) \sum_{j=0}^{\infty} \tilde{\mu}_j^2 \sum_{n=j+m-1}^{\infty} n^{-3} \\
&= O\left(\sum \frac{\tilde{\mu}_j^2}{j^2}\right) < \infty.
\end{aligned}$$

To establish the summability of $\tilde{\mu}_n^2/n^2$, we recall from (4.9) and (4.2) that

$$\begin{aligned}
\tilde{\mu}_n &= O(n^{\alpha-1}) + t_n - \frac{1}{H_m - 1}(n+1) \sum_{k=n}^{\infty} \frac{\hat{t}_k}{(k+1)(k+2)} \\
&\quad + \sum_{j=2}^{m-1} O\left(n^{\alpha_j - 1} \sum_{k=0}^{n-1} \frac{|\hat{t}_k|}{(k+1)^{\alpha_j}}\right),
\end{aligned} \tag{5.21}$$

writing $\alpha_j := \operatorname{Re}\lambda_j$ (with $\alpha = \alpha_2 < 3/2$, since $m \le 26$). Using (5.19), we need only establish the summability of n^{-2} times the square of each of the four terms on the right in (5.21). The first of these verifications is trivial, and the second was carried out at (5.5). For the third we apply the Cauchy–Schwarz inequality [compare (5.6)] to give

$$\begin{aligned}\left[\sum_{k=n}^{\infty}\frac{\hat{t}_k}{(k+1)(k+2)}\right]^2 &\le \frac{1}{n}\left[\sum_{k=n}^{\infty}\sqrt{n}\frac{\sqrt{k}}{(k+1)(k+2)}\frac{|t_k|}{\sqrt{k}}\right]^2 \\ &= O\left(\frac{1}{n}\sum_{k=n}^{\infty}\sqrt{n}k^{-3/2}\frac{t_k^2}{k}\right) = O\left(n^{-1/2}\sum_{k=n}^{\infty}k^{-5/2}t_k^2\right),\end{aligned}$$

whence

$$\begin{aligned}\sum_n\left[\sum_{k=n}^{\infty}\frac{\hat{k}_k}{(k+1)(k+2)}\right]^2 &= O\left(\sum_n n^{-1/2}\sum_{k=n}^{\infty}k^{-5/2}t_k^2\right) \\ &= O\left(\sum_k k^{-5/2}t_k^2k^{1/2}\right) = O\left(\sum_k\frac{t_k^2}{k^2}\right) < \infty\end{aligned}$$

by (5.5) again.

We pause here to note that when $m = 2$ the proof is finished here, and that up to now we have used only (5.5), not the stronger assumption (5.1)(b).

For our fourth and final verification, it suffices [again by invoking (5.19)] to establish the summability of

$$n^{2\rho-4}\left[\sum_{k=1}^{n-1}\frac{|t_k|}{k^\rho}\right]^2 \tag{5.22}$$

for any real $\rho < 3/2$. To do this, we break the sum into $\sum_{k<n^\delta}$ and $\sum_{n^\delta\le k<n}$ and once again invoke (5.19). In the range $\sum_{k<n^\delta}$ we simply use $t_k = O\left(\sqrt{k}\right)$ and note

$$n^{2\rho-4}\left[\sum_{k<n^\delta}O\left(k^{(1/2)-\rho}\right)\right]^2 = O\left(n^{2\rho-4}\left(n^\delta\right)^{3-2\rho}\right) = O\left(n^\tau\right)$$

with $\tau < -1$. In the range $\sum_{n^\delta\le k<n}$ we use Cauchy–Schwarz again:

$$\begin{aligned}n^{2\rho-4}\left[\sum_{n^\delta\le k<n}\frac{|t_k|}{k^\rho}\right]^2 &= n^{2\rho-4}n^{3-2\rho}\left[\sum_{n^\delta\le k<n}\frac{k^{(1/2)-\rho}}{n^{(3/2)-\rho}}\frac{|t_k|}{k^{1/2}}\right]^2 \\ &= O\left(n^{-1}\sum_{n^\delta\le k<n}\frac{k^{(1/2)-\rho}}{n^{(3/2)-\rho}}\frac{t_k^2}{k}\right) = O\left(n^{-1}\max_{n^\delta\le k<n}\frac{t_k^2}{k}\right),\end{aligned}$$

which is summable by assumption (5.1)(b). $\square$

Now we proceed to our "borderline small" case.

Proof of CLT II (Theorem 5.4). Again we use the method of moments together with asymptotic transfer results. Given the similarity to the proof of CLT I, we will be brief here.

Equation (5.9) simply repeats the asymptotic transfer result (4.6). As before, (σ_n^2) satisfies the recurrence (2.3) with (b_n) replaced by (r_n) of (5.17)–(5.18), where again $\mu := K_1/(H_m - 1)$ and $r_j := 0$ for $0 \le j \le m-2$. Here the proofs diverge somewhat. The analogue of Lemma 5.6 is Lemma 5.7 below. Then the asymptotic variance assertions of CLT II follow immediately from Theorem 4.5(c).

If $\sum^\infty \frac{L^2(k)}{k} < \infty$, then from (4.10) applied to L^2 it follows that (5.12) satisfies (5.13)–(5.14) for $k = 1$. Then higher moments are treated exactly as in the proof of CLT I to complete the proof of CLT II.

If $\sum^\infty \frac{L^2(k)}{k} = \infty$, then one uses (4.10), Theorem 4.5(d), and induction to show that the moments (5.12) satisfy

$$\tilde{\mu}_n(2k) \sim \frac{(2k)!}{2^k k!} s^{2k}(n), \qquad k \ge 1 \tag{5.23}$$

and

$$\tilde{\mu}_n(2k-1) = o\left(s^{2k-1}(n)\right), \qquad k \ge 1 \tag{5.24}$$

and thereby complete the proof of CLT II. We omit the details. □

The following cousin to Lemma 5.6 was used in the proof of Theorem 5.4.

Lemma 5.7. *In the context of CLT II, the sequence (r_n) defined for $n \ge m-1$ by*

$$r_n := \frac{1}{\binom{n}{m-1}} \sum_{\mathbf{j}} \left[t_n + \tilde{\mu}_{j_1} + \cdots + \tilde{\mu}_{j_m} - \tilde{\mu}_n\right]^2$$

satisfies

$$r_n \sim (H_m - 1)\sigma^2 n L^2(n).$$

Proof. By (5.9), with $\theta := (3/2)^{\overline{m-1}}/(m! - (3/2)^{\overline{m-1}})$, we have

$$\begin{aligned} r_n &\sim \frac{1}{\binom{n}{m-1}} \sum_{\mathbf{j}} \left[n^{1/2}L(n) - \theta j_1^{1/2} L(j_1) - \cdots - \theta j_m^{1/2} L(j_m) + \theta n^{1/2} L(n)\right]^2 \\ &\sim nL^2(n)\,(m-1)! \int \left[(1+\theta) - \theta \sum_{i=1}^{m-1} x_i^{1/2} - \theta\left(1 - \sum_{i=1}^{m-1} x_i\right)^{1/2}\right]^2 dx_1 \cdots dx_{m-1}, \end{aligned}$$

where the integral, call it J, is over $(x_1, \ldots, x_{m-1}) \in [0,1]^{m-1}$ with sum not exceeding unity. To complete the proof we need only show $J = (H_m - 1)\sigma^2/(m-1)!$, with σ^2 defined at (5.11).

Indeed,

$$\begin{aligned}
J &= \int \Bigg[(1+\theta)^2 + \theta^2 \sum_{i=1}^{m-1} x_i + \theta^2 \left(1 - \sum_{i=1}^{m-1} x_i\right) \\
&\qquad - 2\theta(1+\theta) \sum_{i=1}^{m-1} x_i^{1/2} - 2\theta(1+\theta)\left(1 - \sum_{i=1}^{m-1} x_i\right)^{1/2} \\
&\qquad + \theta^2 \sum_{i,j:\, i \neq j} x_i^{1/2} x_j^{1/2} + 2\theta^2 \left(1 - \sum_{i=1}^{m-1} x_i\right)^{1/2} \sum_{i=1}^{m-1} x_i^{1/2} \Bigg] \, dx_1 \cdots dx_{m-1} \\
&= \left[(1+\theta)^2 + \theta^2\right] \frac{1}{(m-1)!} - 2\theta(1+\theta) m \frac{\Gamma(3/2)}{\Gamma(m+(1/2))} + \theta^2 m(m-1) \frac{[\Gamma(3/2)]^2}{\Gamma(m+1)} \\
&= \left[(1+\theta)^2 + \theta^2\right] \frac{1}{(m-1)!} - 2\theta(1+\theta) m \frac{1}{(3/2)^{\overline{m-1}}} + \theta^2 \frac{\pi/4}{(m-2)!}.
\end{aligned}$$

Plugging in the value of θ and simplifying, we obtain $J = (H_m - 1)\sigma^2/(m-1)!$, as desired.

□

5.3 Periodicity for $m \geq 27$

If $t_n = o(\sqrt{n})$ as in CLT I but $m \geq 27$, then the remainder term $\tilde{\mu}_n := \mu_n - \mu(n+1)$ for the mean—which by (5.2) was $o(\sqrt{n})$ when $m \leq 26$—now satisfies, by the ETT and (4.3) [compare (4.9)]

$$\begin{aligned}
\tilde{\mu}_n &= c_2 \frac{n^{\lambda_2 - 1}}{\Gamma(\lambda_2)} + c_3 \frac{n^{\lambda_3 - 1}}{\Gamma(\lambda_3)} \\
&\quad + m! \sum_{j=2}^{m-1} \frac{1}{\psi'(\lambda_j)} [z^n] \left((1-z)^{-\lambda_j} \int_0^z \hat{T}(\zeta)(1-\zeta)^{\lambda_j - 1} \, d\zeta \right) \\
&\quad + o\left(\sqrt{n}\right) + O\left(n^{\operatorname{Re}\lambda_4 - 1}\right). \qquad (5.25)
\end{aligned}$$

Typically this will lead to the negative result that $(\tilde{\mu}_n)$ [and hence also (r_n) and (σ_n^2)] suffers from periodicity and that there is no natural distributional limit for normalized X_n. Examples are the space requirement functional studied by Chern and Hwang [5] (and others before them [18]) and the shape functional [9]. For recent developments in the case $m > 26$ for the space requirement see [4].

But it is perhaps difficult to establish a *general* negative result, due to cancellations. For example, suppose $T(z)$ equals $(1-z)^{-1}$, so that $t_n \equiv 1$, as studied by Chern and Hwang [5], except perhaps that the initial values $t_0, \ldots, t_{m-2}$ are changed. Then

$$T^{(m-1)}(z) \equiv (m-1)!(1-z)^{-m},$$

whence

$$\begin{aligned} A(z) &= \sum_{n=1}^{\infty} \mu_n z^n = \sum_{j=1}^{m-1} c_j (1-z)^{-\lambda_j} + (m-1)! \sum_{j=1}^{m-1} \frac{(1-z)^{-\lambda_j}}{\psi'(\lambda_j)} \int_0^z (1-\zeta)^{\lambda_j - 2}\, d\zeta \\ &= \sum_{j=1}^{m-1} c_j (1-z)^{-\lambda_j} - \frac{1}{m-1}(1-z)^{-1} - (m-1)! \sum_{j=1}^{m-1} \frac{(1-z)^{-\lambda_j}}{(1-\lambda_j)\psi'(\lambda_j)}. \end{aligned}$$

Now it is possible to choose $t_0, \ldots, t_{m-2}$ so that

$$c_j = \frac{(m-1)!}{(1-\lambda_j)\psi'(\lambda_j)}, \qquad j = 1, \ldots, m-1. \tag{5.26}$$

In that case $A(z) = -\frac{1}{m-1}(1-z)^{-c}$, whence $\tilde{\mu}_n \equiv \mu_n \equiv -1/(m-1)$ [and we see that the chosen values of $t_0, \ldots, t_{m-2}$ are all $-1/(m-1)$], and we get linear variance and asymptotic normality, just as in CLT I, for *every* $2 \le m < \infty$.

One might object that the above example is artificial, in that the toll sequence changes sign. But the same calculation show that if the toll sequence is chosen as above ($t_n \equiv 1$) but with inital values

$$t_j := K(j+1) - \tfrac{1}{m-1}, \qquad 0 \le j \le m-2,$$

then still, for every $m \ge 2$, the sequence $(\tilde{\mu}_n)$ is constant, the variance is linear, and we have asymptotic normality. Further, (t_n) is nonnegative provided $K \ge 1/(m-1)$. [We remark in passing that the choice $K = 1/(m-1)$ leads to the degenerate case of Remark 5.2.] The sequence (t_n) is also nondecreasing (as in most real examples) provided $K \le m/(m-1)^2$.

6 Moderate and large toll functions

In order to describe the limiting distribution of X_n for moderate and large tolls, we will introduce a family of random variables $Y \equiv Y(\beta)$ defined for $\beta > 1/2$, $\beta \ne 1$. Anticipating Lemma 6.1, we need to consider the distributional equation

$$Y \stackrel{\mathcal{L}}{=} \sum_{j=1}^{m} S_j^{\beta} Y_j + 1 \tag{6.1}$$

Here $(Y_j)_{j=1}^m$ are independent copies of Y and $(S_1, \ldots, S_m)$ is uniformly distributed on the $(m-1)$-simplex, independent of $(Y_j)_{j=1}^m$. Recall that the $(m-1)$-simplex is the set

$$\{(s_1, \ldots, s_m) \colon s_j \ge 0 \text{ for } 1 \le j \le m \text{ and } \mathbf{s}_+ = 1\},$$

where $\mathbf{s}_+$ denotes $\sum_{j=1}^m s_j$.

Let $U_{(1)}, \ldots, U_{(m-1)}$ be the order statistics of a sample of size $m-1$ from the uniform distribution on $(0,1)$. They have joint density

$$f_{U_{(1)},\ldots,U_{(m-1)}}(x_1, \ldots, x_{m-1}) \equiv (m-1)! \mathbf{1}(0 < x_1 < \cdots < x_{m-1} < 1)$$

with respect to Lebesgue measure on $\mathbb{R}^{m-1}$, where $\mathbf{1}(A)$ is the indicator of A.

By a change of variables, we find that the joint distribution of the spacings $S_1, \dots, S_m$, defined, with $U_{(0)} := 0$ and $U_{(m)} := 1$, by

$$S_i := U_{(i)} - U_{(i-1)}, \; i = 1, \dots, m,$$

is uniform over the $(m-1)$-simplex:

$$f_{S_1,\dots,S_{m-1}}(s_1, \dots, s_{m-1}) \equiv (m-1)!\mathbf{1}(s_j > 0, j = 1, \dots, m-1; \; \mathbf{s}_+ < 1)\,.$$

When $r_j > -1$ for $1 \le j \le m$, observe that

$$\begin{aligned}
\mathbf{E}\left[\prod_{j=1}^{m} S_j^{r_j}\right] &= (m-1)! \int_{\substack{s_1,\dots,s_{m-1}>0 \\ \mathbf{s}_+<1}} s_1^{r_1} \cdots s_{m-1}^{r_{m-1}} (1-\mathbf{s}_+)^{r_m} \, ds_{m-1} \cdots ds_1 \\
&=: (m-1)! B(r_1+1, \dots, r_m+1) \quad \text{(defining } B \text{ as the integral)} \qquad (6.2) \\
&= (m-1)! \frac{\prod_{j=1}^{m} \Gamma(r_j+1)}{\Gamma(r_1 + \cdots + r_m + m)}.
\end{aligned}$$

Lemma 6.1. *Fix $\beta > 1/2$ with $\beta \neq 1$. Then there exists a unique distribution $\mathcal{L}(Y) \equiv \mathcal{L}(Y(\beta))$ with finite second moment satisfying the distributional identity* (6.1)*.*

Proof. We first observe the that mean of any such distribution is determined by (6.1). Indeed, by taking expectations in (6.1) and using (6.2), we get

$$\mu := \mathbf{E}\, Y = \left(1 - \frac{m!\Gamma(\beta+1)}{\Gamma(\beta+m)}\right)^{-1}$$

since $\beta \neq 1$. Thus we can equivalently consider the distributional identity

$$W \stackrel{\mathcal{L}}{=} \sum_{j=1}^{m} S_j^{\beta} W_j + H,$$

where

$$H := 1 - \mu + \mu \sum_{j=1}^{m} S_j^{\beta}.$$

Here W is restricted to have mean 0 and finite second moment, $(W_j)_{j=1}^m$ are independent copies of W, and $(S_1, \dots, S_m)$ is uniformly distributed on the $(m-1)$-simplex, independent of $(W_j)_{j=1}^m$.

We now employ a standard contraction-method argument [23, 24]. Let d_2 denote the metric on $\mathcal{M}_2(0)$, the space of probability distributions with mean 0 and finite variance, defined by

$$d_2(G_1, G_2) := \min \|X_2 - X_1\|_2,$$

taking the minimum over all pairs of random variables X_1 and X_2 defined on a common probability space with $\mathcal{L}(X_1) = G_1$ and $\mathcal{L}(X_2) = G_2$. Here $\|\cdot\|_2$ denotes L_2-norm.

Let T be the map

$$T : \mathcal{M}_2(0) \to \mathcal{M}_2(0), \quad G \mapsto \mathcal{L}\left(\sum_{j=1}^{m} S_j^{\beta} X_j + H\right),$$

where $(X_j)_{j=1}^m$ are independent with $\mathcal{L}(X_j) = G$, $j = 1, \ldots, m$, and $(S_1, \ldots, S_m)$ is uniformly distributed on the $(m-1)$-simplex, independent of $(X_j)_{j=1}^m$. We show that T is a contraction on $\mathcal{M}_2(0)$; more precisely, that there exists a $\rho < 1$ such that

$$d_2(T(\mathcal{L}(A)), T(\mathcal{L}(B))) \leq \rho d_2(\mathcal{L}(A), \mathcal{L}(B))$$

for all pairs $\mathcal{L}(A)$ and $\mathcal{L}(B)$ in $\mathcal{M}_2(0)$. To bound $d_2(T(\mathcal{L}(A)), T(\mathcal{L}(B)))$, we couple $T(\mathcal{L}(A))$ and $T(\mathcal{L}(B))$ by taking m independent copies (A_j, B_j) of the d_2-optimally coupled (A, B), an independent $(S_1, \ldots, S_m)$, and defining

$$A' := \sum_{j=1}^{m} S_j^{\beta} A_j + H \sim T(\mathcal{L}(A)), \quad B' := \sum_{j=1}^{m} S_j^{\beta} B_j + H \sim T(\mathcal{L}(B)).$$

Now, defining $\mathbf{S} := (S_1, \ldots, S_m)$ and using the law of total variance,

$$\begin{aligned}
&d_2(\mathcal{L}(T(A)), \mathcal{L}(T(B)))^2 \\
&\leq \|B' - A'\|_2^2 = \left\| \sum_{j=1}^{m} S_j^{\beta}(B_j - A_j) \right\|_2^2 = \mathbf{Var}\left[\sum_{j=1}^{m} S_j^{\beta}(B_j - A_j)\right] \\
&= \mathbf{E}\,\mathbf{Var}\left[\sum_{j=1}^{m} S_j^{\beta}(B_j - A_j) \,\middle|\, \mathbf{S}\right] + \mathbf{Var}\,\mathbf{E}\left[\sum_{j=1}^{m} S_j^{\beta}(B_j - A_j) \,\middle|\, \mathbf{S}\right] \\
&= \sum_{j=1}^{m} (\mathbf{E} S_j^{2\beta}) \mathbf{Var}\,[B_j - A_j] = d_2(\mathcal{L}(A), \mathcal{L}(B))^2 \sum_{j=1}^{m} \mathbf{E} S_j^{2\beta} = m! \frac{\Gamma(2\beta+1)}{\Gamma(2\beta+m)} d_2(\mathcal{L}(A), \mathcal{L}(B))^2.
\end{aligned}$$

We need only verify that

$$\rho^2 := m! \frac{\Gamma(2\beta+1)}{\Gamma(2\beta+m)} = \frac{m!}{(2\beta+m-1)\cdots(2\beta+1)} < 1,$$

which is true when $\beta > 1/2$. The existence and uniqueness of $\mathcal{L}(Y)$ now follows from the Banach fixed point theorem [24, Theorem 2]. $\square$

6.1 Moderate toll functions

In the case of moderate toll functions, convergence in distribution and convergence of all moments can be stated as

Theorem 6.2. *If the toll sequence (t_n) satisfies*

$$t_n \sim n^{\beta} L(n) \text{ with } 1/2 < \beta < 1,$$

where L is a slowly varying function and $\alpha < 1+\beta$, then the mean of the corresponding additive functional X_n on m-ary search trees with the random permutation model satisfies

$$\mu_n = \mu n - \frac{(1+\beta)^{\overline{m-1}}}{m! - (1+\beta)^{\overline{m-1}}} n^\beta L(n) + o(n^\beta L(n)), \qquad \mu := \frac{K_1}{H_m - 1}, \tag{6.3}$$

with K_1 defined at (2.12)*. Moreover,*

$$\frac{X_n - \mu n}{n^\beta L(n)} \xrightarrow{\mathcal{L}} Y_\beta,$$

with convergence of all moments.

Remark 6.3. It is well known that $\alpha < 3/2$ for $m \le 26$. In Theorem B.7 we show that α increases with m. Thus for a fixed $\beta \in (1/2, 1)$, the condition $\alpha < 1+\beta$ is equivalent to $m \le m_0$ for some $m_0 \ge 26$.

Proof of Theorem 6.2. We use the notation introduced in the proof of Theorem 5.1 in Section 5.2. Equation (6.3) is simply a restatement of the asymptotic transfer result (4.6).

We show that the moments $\tilde{\mu}_n(k)$ satisfy

$$\tilde{\mu}_n(k) = g_k n^{k\beta} L^k(n) + o(n^{k\beta} L^k(n)) \text{ as } n \to \infty. \tag{6.4}$$

The claim holds for $k = 1$ with

$$g_1 := -\frac{(1+\beta)^{\overline{m-1}}}{m! - (1+\beta)^{\overline{m-1}}} = \left(1 - \frac{m!\Gamma(\beta+1)}{\Gamma(\beta+m)}\right)^{-1}. \tag{6.5}$$

Using (5.16), by induction we get, for $k \ge 2$,

$$\begin{aligned}
r_n(k) = o(n^{k\beta} L^k(n)) & \\
&+ \sum_{\mathbf{k}}^{*} \binom{k}{k_1, \dots, k_m, k_{m+1}} (n^\beta L(n))^{k_{m+1}} \\
&\times \frac{1}{\binom{n}{m-1}} \sum_{\mathbf{j}} g_{k_1} (j_1^\beta L(j_1))^{k_1} \cdots g_{k_m} (j_m^\beta L(j_m))^{k_m} \\
= o(n^{k\beta} L^k(n)) & \\
&+ \sum_{\mathbf{k}}^{*} \binom{k}{k_1, \dots, k_m, k_{m+1}} (n^\beta L(n))^{k} g_{k_1} \cdots g_{k_m} \\
&\times \frac{1}{\binom{n}{m-1}} \sum_{\mathbf{j}} \left(\frac{j_1}{n}\right)^{k_1\beta} \cdots \left(\frac{j_m}{n}\right)^{k_m\beta} \frac{L^{k_1}(j_1) \cdots L^{k_m}(j_m)}{L^{k_1+\cdots+k_m}(n)}.
\end{aligned}$$

But [recall the definition of B at (6.2)]

$$\frac{1}{\binom{n}{m-1}} \sum_{\mathbf{j}} \left(\frac{j_1}{n}\right)^{k_1\beta} \cdots \left(\frac{j_m}{n}\right)^{k_m\beta} \frac{L^{k_1}(j_1) \cdots L^{k_m}(j_m)}{L^{k_1+\cdots+k_m}(n)} \to (m-1)! B(k_1\beta+1, \dots k_m\beta+1)$$

so that

$$\begin{aligned}r_n(k) &= o(n^{k\beta}L^k(n))\\ &\quad + n^{k\beta}L^k(n)(m-1)!\sum_{\mathbf{k}}{}^{*}\binom{k}{k_1,\dots,k_m,k_{m+1}}g_{k_1}\cdots g_{k_m}B(k_1\beta+1,\dots,k_m\beta+1).\end{aligned}$$

Using Theorem 4.5, with $v = k\beta > 1$, we get

$$\begin{aligned}&\tilde{\mu}_n(k) = o(n^{k\beta}L^k(n))\\ &+n^{k\beta}L^k(n)\frac{(m-1)!}{1-\frac{m!\Gamma(k\beta+1)}{\Gamma(k\beta+m)}}\sum_{\mathbf{k}}{}^{*}\binom{k}{k_1,\dots,k_m,k_{m+1}}g_{k_1}\cdots g_{k_m}B(k_1\beta+1,\dots,k_m\beta+1).\end{aligned}$$

Thus, defining g_k recursively as

$$g_k = \frac{(m-1)!}{1-\frac{m!\Gamma(k\beta+1)}{\Gamma(k\beta+m)}}\sum_{\mathbf{k}}{}^{*}\binom{k}{k_1,\dots,k_{m+1}}g_{k_1}\cdots g_{k_m}B(k_1\beta+1,\dots,k_m\beta+1), \tag{6.6}$$

with $g_0 = 1$, we see that (6.4) holds for all $k \geq 0$.

By Lemma 6.4 (to follow) and the method of moments (cf., e.g., [7, Sections 4.4 and 4.5]), the g_k's are the moments of a uniquely determined distribution, say $\mathcal{L}(\widehat{Y})$, and

$$\frac{X_n - \mu n}{n^{\beta}L(n)} \xrightarrow{\mathcal{L}} \widehat{Y}$$

with convergence of all moments. It remains to show that $\widehat{Y} \stackrel{\mathcal{L}}{=} Y(\beta)$.

Define

$$\widetilde{Y} := \sum_{j=1}^{m} S_j^{\beta}\widehat{Y}_j + 1, \tag{6.7}$$

where $(\widehat{Y}_j)_{j=1}^m$ are independent copies of $\widehat{Y}$ and $(S_1,\dots,S_m)$ is uniformly distributed on the $(m-1)$-simplex, independent of $(\widehat{Y}_j)_{j=1}^m$. We will show that $\widehat{Y} \stackrel{\mathcal{L}}{=} \widetilde{Y}$, and then, by (6.7), $\mathcal{L}(\widehat{Y})$ satisfies the distributional identity (6.1) and has finite second moment. By Lemma 6.1, $\widehat{Y} \stackrel{\mathcal{L}}{=} Y(\beta)$, as desired.

To show $\widehat{Y} \stackrel{\mathcal{L}}{=} \widetilde{Y}$, it suffices to show that $\widehat{Y}$ and $\widetilde{Y}$ have the same moments. Letting $\sum_{\mathbf{k}}$ denote (as before) the sum over $(m+1)$-tuples $(k_1,\dots,k_{m+1})$ of nonnegative integers

summing to k, and using (6.7), (6.2), and (6.6),

$$
\begin{aligned}
\mathbf{E}\,\widetilde{Y}^k &= \sum_{\mathbf{k}} \binom{k}{k_1,\dots,k_{m+1}} \mathbf{E}\left[\prod_{j=1}^m (S_j^\beta \widehat{Y}_j)^{k_j}\right] \\
&= \sum_{\mathbf{k}} \binom{k}{k_1,\dots,k_{m+1}} (m-1)!B(k_1\beta+1,\dots k_m\beta+1) g_{k_1}\cdots g_{k_m} \\
&= (m-1)! \sum_{\mathbf{k}}^{*} \binom{k}{k_1,\dots,k_{m+1}} B(k_1\beta+1,\dots k_m\beta+1) g_{k_1}\cdots g_{k_m} \\
&\quad + m!B(k\beta+1,1,\dots,1)g_k \\
&= \left[1 - \frac{m!\Gamma(k\beta+1)}{\Gamma(k\beta+m)}\right] g_k + \frac{m!\Gamma(k\beta+1)}{\Gamma(k\beta+m)} g_k = g_k = \mathbf{E}\,\widehat{Y}^k,
\end{aligned}
$$

where (as before) $\sum_{\mathbf{k}}^{*}$ denotes the same sum as $\sum_{\mathbf{k}}$ with the additional restriction that $k_i < k$ for $i = 1,\dots,m$. □

Lemma 6.4. *The moments (g_k) uniquely determine the distribution $\mathcal{L}(Y)$.*

Proof. Define $\gamma_k := g_k/k!$. It suffices to show (by Carleman's condition) that there exists an M such that $\gamma_k \le M^k$ for all k sufficiently large. We proceed by induction. Indeed, by (6.6) we know

$$
\begin{aligned}
\gamma_k &= \frac{(m-1)!}{1 - \frac{m!\Gamma(k\beta+1)}{\Gamma(k\beta+m)}} \sum_{\mathbf{k}}^{*} \frac{1}{k_{m+1}!}\left(\prod_{j=1}^m \gamma_{k_j}\right) B(k_1\beta+1,\dots,k_m\beta+1) \\
&\le M^k \frac{(m-1)!}{1 - \frac{m!\Gamma(k\beta+1)}{\Gamma(k\beta+m)}} \sum_{\mathbf{k}}^{*} \frac{M^{-k_{m+1}}}{k_{m+1}!} B(k_1\beta+1,\dots,k_m\beta+1)
\end{aligned}
$$

by the induction hypothesis. So it is certainly sufficient to show that

$$
\begin{aligned}
&\sum_{\mathbf{k}}^{*} \frac{M^{-k_{m+1}}}{k_{m+1}!} B(k_1\beta+1,\dots,k_m\beta+1) \\
&= \sum_{k_{m+1}=0}^{k} \frac{M^{-k_{m+1}}}{k_{m+1}!} \Gamma((k-k_{m+1})\beta+m)^{-1} \sum_{\substack{0\le k_1,\dots,k_m<k \\ k_1+\cdots+k_m=k-k_{m+1}}} \prod_{j=1}^m \Gamma(k_j\beta+1) \\
&\to 0 \quad \text{as } k\to\infty.
\end{aligned} \tag{6.8}
$$

For this, fix a value of $k_{m+1} \in \{0,1,2,\dots\}$, and consider the sum

$$
\sum_{\substack{0\le k_1,\dots,k_m<k \\ k_1+\cdots+k_m=k-k_{m+1}}} \prod_{j=1}^m \Gamma(k_j\beta+1). \tag{6.9}
$$

By log-convexity of Γ [1, 6.4.1], taking I to be $(0,\infty)$ and g to be $\log\Gamma$ in Proposition 3.C.1 of [20], the logarithm of the product in (6.9) is Schur-convex on $(0,\infty)^m$.

Thus applying Proposition 5.C.2 of [20] with $m = 0$ there, the biggest terms in the sum correspond to $k_j = k - k_{m+1}$ for j equal to some j_0 and $k_j = 0$ otherwise; together, these m terms contribute $m\Gamma((k - k_{m+1})\beta + 1)$ to the sum. If $k > k_{m+1}$, there are other terms, the biggest of which corresponds to having one of the k_j's be $k - k_{m+1} - 1$, one be 1, and the rest be 0. (This follows from Proposition 5.C.1 of [20] with $m = 0$ and $M = k - k_{m+1} - 1$ there.) The total number of terms in the sum (6.9) is at most $\binom{m-1+(k-k_{m+1})}{m-1}$. So the remaining contribution to (6.9) is at most

$$\binom{m-1+(k-k_{m+1})}{m-1}\Gamma((k - k_{m+1} - 1)\beta + 1)\Gamma(\beta + 1).$$

We have found that the left side of (6.8) is bounded by

$$\sum_{k_{m+1}=0}^{\infty} \frac{M^{-k_{m+1}}}{k_{m+1}!}\mathbf{1}(k_{m+1} \le k)\, f(k - k_{m+1}),$$

where

$$f(k) \le \frac{m}{(k\beta+1)\cdots(k\beta+m-1)} + \Gamma(\beta+1)\frac{\binom{m-1+k}{m-1}\Gamma((k-1)\beta+1)}{\Gamma(k\beta+m)} \tag{6.10}$$

$$\le \frac{\binom{m-1+k}{m-1}\Gamma(k\beta+1)}{\Gamma(k\beta+m)} = \frac{1}{(m-1)!}\frac{(k+1)\cdots(k+(m-1))}{(k\beta+1)\cdots(k\beta+(m-1))}, \tag{6.11}$$

which is a bounded function of k. To apply the dominated convergence theorem, it suffices to show that the right side of (6.10) tends to 0 as $k \to \infty$, which follows from Stirling's approximation and the fact that $\beta > 0$. $\square$

When t_n satisfies the conditions in Theorem 6.2 but $\alpha \ge 1+\beta$ then [compare (5.25)],

$$\begin{aligned}\tilde{\mu}_n = {} & c_2\frac{n^{\lambda_2-1}}{\Gamma(\lambda_2)} + c_3\frac{n^{\lambda_3-1}}{\Gamma(\lambda_3)} \\ & + m!\sum_{j=1}^{m-1}\frac{1}{\psi'(\lambda_j)}[z^n]\left((1-z)^{-\lambda_j}\int_0^z \hat{T}(\zeta)(1-\zeta)^{\lambda_j-1}\,d\zeta\right) \\ & + o(n^{\beta}) + O(n^{\mathrm{Re}(\lambda_4-1)}),\end{aligned}$$

and typically this leads to periodicity.

6.2 Large toll functions

If $t_n \sim n^{\beta}L(n)$, where $\beta > 1$ and L is slowly varying function, then we have convergence in distribution for all values of m. We state the result, omitting the proof, as it is very similar to that of Theorem 6.2.

Theorem 6.5. *If the toll sequence (t_n) satisfies*

$$t_n \sim n^{\beta}L(n), \textit{with } \beta > 1,$$

where L is a slowly varying function, then

$$\frac{X_n}{n^\beta L(n)} \xrightarrow{\mathcal{L}} Y_(\beta)$$

with convergence of all moments, where $\mathcal{L}(Y(\beta))$ is the unique distribution satisfying (6.1).

Presumably, the borderline case $t_n \sim nL(n)$ where L is a slowly varying function can also be handled using the techniques of this paper, but we have not pursued this. The specific choice $t_n \equiv n-(m-1)$ for $n \geq m-1$ corresponds to the well-studied total path length of a random m-ary search tree. The corresponding additive functional measures the number of basic operations in m-ary `Quicksort`. As is well known, the number of basic operations has mean $\Theta(n \log n)$ and standard deviation $\Theta(n)$. See [13] for details in this case and [22, Corollary 5.2] for a characterization of the limiting distribution of the path length.

A Solution of an Euler differential equation

We now provide the proof of equation (2.6) in Theorem 2.2, which states the general solution of the differential equation (2.4) with initial conditions $a_j = b_j$, $0 \leq j \leq m-2$. This linear differential equation can be written in the form

$$\mathbf{L}g = h \tag{A.1}$$

where the operator $\mathbf{L}$ is defined as

$$(\mathbf{L}g)(z) := g^{(m-1)}(z) - m!(1-z)^{-(m-1)}g(z). \tag{A.2}$$

We seek the solution $g = A$ corresponding to input $h = B^{(m-1)}$.

Equations of the form (A.1)–(A.2) are members of a class known as *Euler differential equations*. In this appendix we discuss a general method for solving Euler equations, restricting attention, for the sake of definiteness and practicality, to (A.1)–(A.2). We assume that the reader is familiar with the theory of linear differential equations with constant coefficients (see, e.g., [3]).

For brevity we have omitted several routine proofs. A fuller version of Appendices A and B.2 may be found in the technical report [11].

The homogeneous solution

The technique for solving $\mathbf{L}g = 0$ is quite easily summarized: make the change of variable $z = 1 - e^{-x}$, that is, $x = \ln\left((1-z)^{-1}\right)$. For notational convenience we will abbreviate $\ln\left((1-z)^{-1}\right)$ as $L(z)$. Then consider the function $\tilde{g}$ defined by

$$\tilde{g}(x) := g(1-e^{-x}), \quad \text{i.e.,} \quad g(z) = \tilde{g}(L(z)). \tag{A.3}$$

Lemma A.1. *The derivatives of g are related to those of $\tilde{g}$ by*

$$g^{(k)}(z) = (1-z)^{-k} \sum_{j=0}^{k} \begin{bmatrix} k \\ j \end{bmatrix} \tilde{g}^{(j)}(L(z)), \tag{A.4}$$

where $\left[\begin{smallmatrix} k \\ j \end{smallmatrix}\right]$ denotes a signless Stirling number of the first kind.

Proof. The proof is a straightforward induction on k using standard identities for the Stirling numbers [15, § 1.2.6]. □

The left-hand side of (A.1) can hence be expressed as

$$(\mathbf{L}g)(z) = (1-z)^{-(m-1)} \left\{ \sum_{j=0}^{m-1} \begin{bmatrix} m-1 \\ j \end{bmatrix} \tilde{g}^{(j)}(L(z)) - m!\tilde{g}(L(z)) \right\}$$

so that solving $\mathbf{L}g = 0$ is equivalent to solving $\widetilde{\mathbf{L}}\tilde{g} = 0$, where

$$\widetilde{\mathbf{L}}\tilde{g}(x) := \sum_{j=0}^{m-1} \begin{bmatrix} m-1 \\ j \end{bmatrix} \tilde{g}^{(j)}(x) - m!\tilde{g}(x). \tag{A.5}$$

But this is a linear differential equation with constant coefficients. Its *indicial polynomial*, or *characteristic polynomial*, is

$$\psi_m(\lambda) \equiv \psi(\lambda) := \sum_{j=0}^{m-1} \begin{bmatrix} m-1 \\ j \end{bmatrix} \lambda^j - m! = \lambda^{\overline{m-1}} - m!, \tag{A.6}$$

the last equality following from [15, 1.2.6-(44)]. For more on this polynomial see Appendix B. From the discussion in [17] it follows that there are $m-1$ distinct (and nonzero) roots of ψ, call them $\lambda_1, \ldots, \lambda_{m-1}$ arranged in nonincreasing order of real parts. Thus the functions $\exp(\lambda_j x)$ are $m-1$ linearly independent solutions of (A.5) and hence the functions $(1-z)^{-\lambda_j}$ form a basis of solutions to $\mathbf{L}g = 0$.

A particular solution

Lemma A.2. *The particular solution to $\mathbf{L}g = h$ with vanishing initial conditions (through order $m-2$) is*

$$g_p(z) = \sum_{j=1}^{m-1} \frac{(1-z)^{-\lambda_j}}{\psi'(\lambda_j)} \int_0^z h(\zeta)(1-\zeta)^{\lambda_j + m - 2}\, d\zeta.$$

The initial conditions

Having computed a basis of solutions to the homogeneous equation and a particular solution to the inhomogeneous equation, so far we have established equation (2.6) in Theorem 2.2 modulo determination of the coefficients $c_1, \ldots, c_{m-1}$ at (2.9).

The fact that the initial conditions for g_p vanish make it simple to solve $\mathbf{L}g = h$ for specified initial conditions: One need only match up the initial conditions of the homogeneous solutions.

Proposition A.3. *If the general complementary solution to* $\mathbf{L}g = h$ *is written in the form*

$$g_c(z) = \sum_{j=1}^{m-1} A_j (1-z)^{-\lambda_j}, \tag{A.7}$$

then the constants A_j *are given by*

$$A_j = \frac{m!}{\psi'(\lambda_j)} \sum_{k=0}^{m-2} \frac{g_c^{(k)}(0)}{\lambda_j^{\overline{k+1}}}, \quad j = 1, \ldots, m-1.$$

B Properties of the indicial polynomial

The indicial polynomial

$$\psi(\lambda) \equiv \psi_m(\lambda) := \lambda^{\overline{m-1}} - m! \tag{B.1}$$

plays an important role in the analysis of random m-ary search trees. We will enumerate a few useful identities involving the polynomial in this appendix.

It is well known [17, Chapter 3] that ψ_m has $m-1$ distinct roots $2 = \lambda_1, \lambda_2, \ldots, \lambda_{m-1}$ listed in nonincreasing order of real part. As in [17, Chapter 3] we introduce

$$\alpha \equiv \alpha_m := \max_{2 \le j \le m-1} \operatorname{Re}(\lambda_j);$$

that is, α is the second largest real part among all the roots of the indicial polynomial (B.1). We list some important properties of the roots of (B.1) stated in [17, §3.3]:

(i) The number $-m$ is a root if and only if m is odd. All other roots of $\psi(\lambda)$ are simple, non-real roots.

(ii) No two roots have the same real part unless they are mutually conjugate. (This follows from the strict increasingness of $|(s+it)^{\overline{m-1}}|$ in $|t|$.)

B.1 Monotonicity of α in m

One can check easily that for $3 \le m \le 6$, α is strictly increasing in m. We will now prove this fact (Theorem B.7) for all $m \ge 3$. To do so we build upon ideas in Appendix A of [19].

Claim B.1. *For any* $m \ge 6$ *and* $-\infty < x < 2$,

$$g_m(x) := \inf\{y > 0 : (2-x+iy)(3-x+iy)\ldots(m-x+iy) \text{ is positive real}\} \tag{B.2}$$

is positive and finite, and the infimum is achieved.

Proof. If $z = 2 - x + iy$ with $-\infty < x < 2$ and $y \ge 0$, then

$$\sum_{j=2}^{m} \arg(j - x + iy) = \sum_{j=2}^{m} \arctan \frac{y}{j-x}.$$

This last expression is strictly increasing in $y \geq 0$, with value 0 at 0 and (for $m \geq 6$) limit

$$\sum_{j=2}^{m} \frac{\pi}{2} = (m-1)\frac{\pi}{2} \geq \frac{5\pi}{2}$$

as $y \to \infty$. It is therefore clear that $g_m(x)$ is positive and finite and in fact is characterized as the unique root to

$$\sum_{j=2}^{m} \arctan \frac{g_m(x)}{j-x} = 2\pi.$$

The set on the right in (B.2) is discrete, and $g_m(x)$ is its smallest element. $\square$

Claim B.2. *The function g_m in* (B.2) *is (for $m \geq 6$) strictly decreasing in $-\infty < x < 2$.*

Proof. If $-\infty < x_1 < x_2 < 2$, then

$$2\pi = \sum_{j=2}^{m} \arctan \frac{g_m(x_1)}{j-x_1} < \sum_{j=2}^{m} \arctan \frac{g_m(x_1)}{j-x_2},$$

so $g_m(x_2) < g_m(x_1)$. $\square$

Claim B.3. *The function g_m in* (B.2) *is (for $m \geq 6$) continuous in $-\infty < x < 2$.*

Proof. Let $-\infty < x_1 < x_2 < 2$. In Claim B.2 we have seen $g_m(x_2) < g_m(x_1)$. To complete the proof of Claim B.3, we will show that

$$g_m(x_2) \geq \frac{2-x_2}{2-x_1} g_m(x_1). \tag{B.3}$$

Indeed,

$$2\pi = \sum_{j=2}^{m} \arctan \frac{g_m(x_1)}{j-x_1} = \sum_{j=2}^{m} \arctan \left[\frac{j-x_2}{j-x_1} \times \frac{g_m(x_1)}{j-x_2} \right] \geq \sum_{j=2}^{m} \arctan \frac{\frac{2-x_2}{2-x_1} g_m(x_1)}{j-x_2},$$

so (B.3) follows. $\square$

Corollary B.4. *For any $m \geq 6$ and $-\infty < x < 2$, define*

$$f_m(x) := (2 - x + i g_m(x))(3 - x + i g_m(x)) \ldots (m - x + i g_m(x)),$$

which by Claim B.1 is positive real-valued. Then f_m is continuous.

Proof. This is immediate from Claim B.3. $\square$

Claim B.5. *For fixed $-\infty < x < 2$, $g_m(x)$ is strictly decreasing in $m \geq 6$.*

Proof. If $m \geq 6$, then

$$2\pi = \sum_{j=2}^{m} \arctan \frac{g_m(x)}{j-x} < \sum_{j=2}^{m+1} \arctan \frac{g_m(x)}{j-x},$$

so $g_{m+1}(x) < g_m(x)$. □

Lemma B.6. *For $m \geq 4$, let $\lambda_m = \alpha_m + i\beta_m$ (with $\beta_m > 0$) be a root of the indicial polynomial ψ_m with second largest real part. Then $|\lambda_m + m - 1| < m + 1$.*

Proof. If $|\lambda_m + m - 1| \geq m + 1$, then by the triangle inequality $|\lambda_m + j| > j + 2$ for all $j = 0, \ldots, m-2$. But then $|\lambda_m^{\overline{m-1}}| = \prod_{j=0}^{m-2} |\lambda_m + j| > m!$, which is a contradiction. □

Combining our preliminary results we can now prove the asserted monotonicity of α.

Theorem B.7. *The second largest real part, α_m, among roots of the indicial polynomial ψ_m is strictly increasing in $m \geq 3$.*

Proof. For $m \leq 6$ the result is easily verified. We now show that $\alpha_{m+1} > \alpha_m$ for $m \geq 6$. Observe $f_{m+1}(0) = |f_{m+1}(0)| > |(m+1)!| = (m+1)!$ and

$$\begin{aligned}
f_{m+1}(2-\alpha_m) &= |f_{m+1}(2-\alpha_m)| \\
&= \big|[\alpha_m + ig_{m+1}(2-\alpha_m)] \ldots [\alpha_m + (m-2) + ig_{m+1}(2-\alpha_m)] \\
&\qquad \times [\alpha + (m-1) + ig_{m+1}(2-\alpha_m)]\big| \\
&< \big|[\alpha_m + ig_m(2-\alpha_m)] \ldots [\alpha_m + (m-2) + ig_m(2-\alpha_m)] \qquad \text{(B.4)} \\
&\qquad \times [\alpha_m + (m-1) + ig_m(2-\alpha_m)]\big| \\
&\leq m!|\alpha_m + (m-1) + i\beta_m| \qquad \text{(B.5)} \\
&= m!|\lambda_m + m - 1| < (m+1)!.
\end{aligned}$$

Inequality (B.4) follows from Claim B.5, inequality (B.5) holds since $g_m(2-\alpha_m) \leq \beta_m$, and the last inequality is a consequence of Lemma B.6. Therefore, by Corollary B.4, $f_{m+1}(x) = (m+1)!$ for some $0 < x < 2 - \alpha_m$, and so $\lambda^{\overline{m}} = (m+1)!$ for some λ with $\operatorname{Re}\lambda > \alpha_m$. That is, $\alpha_{m+1} > \alpha_m$, as desired. □

Remark B.8 *(Asymptotics of α_m).* Using (B.7) and (B.8) from [18] and the characterization of $g_m(x)$ in Claim B.1, we can establish

$$\alpha_m = 2 - (1 + o(1))\, 2\pi^2 \left(\frac{\pi^2}{6} - 1\right) \ln^{-3} m,$$
$$\beta_m = (1 + o(1))\, 2\pi \ln^{-1} m;$$

we omit the proof. Thus the Proposition in Appendix B of [18] is asymptotically optimal, to first order.

B.2 Identities involving the indicial polynomial

Identity B.9. When $\lambda \notin \{\lambda_1, \dots, \lambda_{m-1}\}$,

$$\sum_{j=1}^{m-1} \frac{1}{(\lambda - \lambda_j)\psi'(\lambda_j)} = \frac{1}{\psi(\lambda)}$$

For r and n positive integers, let $H_n^{(r)}$ denote the rth-order harmonic number

$$H_n^{(r)} := \sum_{j=1}^{n} \frac{1}{j^r}.$$

When $r = 1$ we will use $H_n := H_n^{(1)}$ for the usual (1st-order) harmonic number.

Identity B.10. For $0 \le k \le m - 3$,

$$\sum_{j=1}^{m-1} \frac{\lambda_j^k}{\psi'(\lambda_j)} = 0.$$

Identity B.11.

$$\psi'(2) = m!(H_m - 1) \qquad \text{and} \qquad \psi''(2) = m![(H_m - 1)^2 - (H_m^{(2)} - 1)].$$

Identity B.12.

$$\sum_{j=2}^{m-1} \frac{1}{(\lambda_j - 2)\psi'(\lambda_j)} = \frac{1}{2(m!)} \left[1 - \frac{H_m^{(2)} - 1}{(H_m - 1)^2} \right].$$

Acknowledgment. The authors thank two anonymous referees for helpful suggestions.

References

[1] M. Abramowitz and I. A. Stegun. *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, volume 55 of *National Bureau of Standards Applied Mathematics Series*. For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C., 1964.

[2] N. H. Bingham, C. M. Goldie, and J. L. Teugels. *Regular variation*, volume 27 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1989.

[3] W. E. Boyce and R. C. DiPrima. *Elementary differential equations and boundary value problems*. John Wiley & Sons, 4th edition, 1986.

[4] B. Chauvin and N. Pouyanne. m-ary search trees when $m > 26$: a strong asymptotics for the space requirements, 2003. Random Structures & Algorithms. To appear. Available at `http://www.math.uvsq.fr/~chauvin/publisenglish.html`.

[5] H.-H. Chern and H.-K. Hwang. Phase changes in random m-ary search trees and generalized quicksort. *Random Structures Algorithms*, 19(3-4):316–358, 2001. Analysis of algorithms (Krynica Morska, 2000).

[6] H.-H. Chern, H.-K. Hwang, and T.-H. Tsai. An asymptotic theory for Cauchy-Euler differential equations with applications to the analysis of algorithms. *J. Algorithms*, 44(1):177–225, 2002. Analysis of algorithms.

[7] K. L. Chung. *A course in probability theory.* Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York-London, second edition, 1974. Probability and Mathematical Statistics, Vol. 21.

[8] R. P. Dobrow and J. A. Fill. Multiway trees of maximum and minimum probability under the random permutation model. *Combin. Probab. Comput.*, 5(4):351–371, 1996.

[9] J. A. Fill. On the distribution of binary search trees under the random permutation model. *Random Structures Algorithms*, 8(1):1–25, 1996.

[10] J. A. Fill, P. Flajolet, and N. Kapur. Singularity analysis, Hadamard products, and Tree recurrences, 2003, arXiv:math.CO/0306225. Submitted for publication.

[11] J. A. Fill and N. Kapur. Transfer theorems and asymptotic distributional results for m-ary search trees, arXiv:math.PR/0306050. Version 1 of the present paper.

[12] P. Flajolet and A. Odlyzko. Singularity analysis of generating functions. *SIAM J. Discrete Math.*, 3(2):216–240, 1990.

[13] P. Hennequin. *Analyse en moyenne d'algorithme, tri rapide et arbres de recherche.* PhD thesis, Ecole Polytechnique, 1991.

[14] H.-K. Hwang and R. Neininger. Phase change of limit laws in the quicksort recurrence under varying toll functions. *SIAM J. Comput.*, 31(6):1687–1722, 2002.

[15] D. E. Knuth. *The art of computer programming. Volume 1.* Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont., 3rd edition, 1997.

[16] D. E. Knuth. *The art of computer programming. Volume 3.* Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont., 2nd edition, 1998.

[17] H. M. Mahmoud. *Evolution of random search trees.* John Wiley & Sons Inc., New York, 1992. A Wiley-Interscience Publication.

[18] H. M. Mahmoud and B. Pittel. Analysis of the space of search trees under the random insertion algorithm. *J. Algorithms*, 10(1):52–75, 1989.

[19] H. M. Mahmoud and R. T. Smythe. Probabilistic analysis of bucket recursive trees. *Theoret. Comput. Sci.*, 144(1-2):221–249, 1995. Special volume on mathematical analysis of algorithms.

[20] A. W. Marshall and I. Olkin. *Inequalities: theory of majorization and its applications*. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1979.

[21] R. R. Muntz and R. C. Uzgalis. Dynamic storage allocation for binary search trees in a two-level memory. In *Proceedings of Princeton Conference on Information Sciences and Systems*, volume 4, pages 345–349, 1971.

[22] R. Neininger and L. Rüschendorf. On the internal path length of d-dimensional quad trees. *Random Structures Algorithms*, 15(1):25–41, 1999.

[23] U. Rösler. On the analysis of stochastic divide and conquer algorithms. *Algorithmica*, 29(1-2):238–261, 2001. Average-case analysis of algorithms (Princeton, NJ, 1998).

[24] U. Rösler and L. Rüschendorf. The contraction method for recursive algorithms. *Algorithmica*, 29(1-2):3–33, 2001. Average-case analysis of algorithms (Princeton, NJ, 1998).